

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 18, 2001, 07:16:03 : Search time 885.31 seconds
(without alignments)
2010.484 Million cell updates/sec

Title: US-09-199-129A-1
Perfect score: 254
Sequence: 1 ggaatcgaagcaccctta.....atgaacaaagcttcagaca 254

Scoring table: IDENTITY-NUC
Gapop 10.0, Capext 1.0

Searched: 7991742 seqs, 3503743858 residues
15983484

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: em_estba:*
42: em_estfun:*
43: em_esthum1:*

44: em_esthum2:*
45: em_esthum3:*
46: em_esthum4:*
47: em_esthum5:*
48: em_esthum6:*
49: em_esthum7:*
50: em_esthum8:*
51: em_esthum9:*
52: em_esthum10:*
53: em_esthum11:*
54: em_esthum12:*
55: em_esthum13:*
56: em_esthum14:*
57: em_esthum15:*
58: em_esthum16:*
59: em_esthum17:*
60: em_esthum18:*
61: em_esthum19:*
62: em_esthum20:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estov1:*
68: em_estov2:*
69: em_estp11:*
70: em_estp12:*
71: em_estp13:*
72: em_estp14:*
73: em_estp15:*
74: em_estro1:*
75: em_estro2:*
76: em_estro3:*
77: em_estro4:*
78: em_estro5:*
79: em_estro6:*
80: em_estro7:*
81: em_estro8:*
82: em_estro9:*
83: em_estro10:*
84: em_estro11:*
85: em_estro12:*
86: em_estro13:*
87: gb_est41:*
88: gb_est42:*
89: gb_est43:*
90: gb_est44:*
91: gb_est45:*
92: gb_est46:*
93: gb_est47:*
94: gb_est48:*
95: gb_est49:*
96: gb_est50:*
97: gb_est51:*
98: gb_est52:*
99: gb_est53:*
100: gb_est54:*
101: gb_est55:*
102: gb_est56:*
103: gb_est57:*
104: gb_est58:*
105: gb_est59:*
106: gb_est60:*
107: gb_est61:*
108: gb_est62:*
109: gb_est63:*
110: gb_est64:*
111: gb_est65:*
112: em_esthum21:*
113: em_esthum22:*
114: em_esthum23:*
115: em_estom1:*
116: em_estom2:*

117: em_estp16:*
 118: em_estp17:*
 119: em_estp18:*
 120: em_estp14:*
 121: em_estp15:*
 122: em_estp16:*
 123: em_estp17:*
 124: em_estp18:*
 125: em_estp19:*
 126: gb_est58:*
 127: gb_est59:*
 128: gb_est60:*
 129: gb_est61:*
 130: gb_est62:*
 131: gb_est63:*
 132: gb_est64:*
 133: gb_est65:*
 134: gb_est66:*
 135: gb_est67:*
 136: gb_est68:*
 137: gb_est69:*
 138: gb_est70:*
 139: gb_est71:*
 140: gb_est72:*
 141: gb_est73:*
 142: gb_est74:*
 143: gb_est75:*
 144: gb_est76:*
 145: gb_est77:*
 146: gb_est78:*
 147: gb_est79:*
 148: gb_est80:*
 149: gb_est81:*
 150: gb_est82:*
 151: gb_est83:*
 152: gb_est84:*
 153: gb_est85:*
 154: gb_est86:*
 155: gb_est87:*
 156: gb_est88:*
 157: gb_est89:*
 158: gb_est90:*
 159: gb_est91:*
 160: gb_est92:*
 161: gb_est93:*
 162: gb_est94:*
 163: gb_est95:*
 164: gb_est96:*
 165: gb_est97:*
 166: gb_est98:*
 167: gb_est99:*
 168: gb_est100:*
 169: gb_est101:*
 170: gb_est102:*
 171: gb_est103:*
 172: gb_est104:*
 173: gb_est105:*
 174: gb_est106:*
 175: gb_est107:*
 176: gb_est108:*
 177: gb_est109:*
 178: gb_est110:*
 179: gb_est111:*
 180: gb_est112:*
 181: gb_est113:*
 182: gb_est114:*
 183: gb_est115:*
 184: gb_est116:*
 185: gb_est117:*
 186: gb_est118:*
 187: gb_est119:*
 188: gb_est120:*
 189: gb_est121:*

190: gb_gss25:*
 191: gb_gss26:*
 192: gb_gss27:*
 193: gb_gss28:*

Print. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	127.6	50.2	477	28	AL369727	AL369727 MBLA32H11
2	127.6	50.2	604	105	BE204104	BE204104 EST396780
3	127.6	50.2	628	105	BE204201	BE204201 EST396877
4	119.4	47.0	948	94	AW722203	AW722203 GA_Ea001
5	110.4	43.5	463	108	BE462089	BE462089 EST324351
6	110.4	43.5	468	137	BE921143	BE921143 EST424912
7	108.8	42.8	263	108	BE449253	BE449253 EST356108
8	107.8	42.4	529	97	AW934135	AW934135 EST359978
9	107.8	42.4	807	108	BE420773	BE420773 HMM002.D0
10	104.6	41.2	714	105	BE193174	BE193174 HVMEM008
11	103.4	40.7	547	110	BE598903	BE598903 P11_83_A0
12	102.8	40.5	290	28	AT002063	AT002063 AT001761
13	102.8	40.5	411	28	AT001761	AT001761 AT001761
14	102.8	40.5	564	93	AW624478	AW624478 EST322423
15	99.6	39.2	646	106	BE320109	BE320109 NF025B04R
16	99.2	39.1	699	92	AW584510	AW584510 N210564e
17	99.2	39.1	509	92	AW560421	AW560421 EST315469
18	98.6	38.8	417	26	A1898214	A1898214 EST267657
19	96.4	38.0	380	144	R65185	R65185 13689 Lambd
20	94.6	37.2	337	111	BE662436	BE662436 EST414417
21	94.2	37.1	191	108	BE459125	BE459125 EST414417
22	94	37.0	555	27	A1993867	A1993867 701515234
23	82.4	32.4	474	29	A0084980	A0084980 A0084980
24	80.6	31.7	494	94	AW754529	AW754529 PC03C03 P
25	78.6	30.9	584	190	AG033012	AG033012 Oryza sat
26	77.2	30.4	337	111	BE662436	BE662436 ST85/ST85
27	77.2	30.4	344	87	AW226329	AW226329 ST85/ST85
28	77.2	30.4	351	90	AW437935	AW437935 ST78E07 P
29	77.2	30.4	380	87	AW226271	AW226271 ST80D09 P
30	77.2	30.4	389	39	AW056809	AW056809 ST56D12 P
31	77.2	30.4	390	39	AW056655	AW056655 ST54C03 P
32	77.2	30.4	460	92	AW587778	AW587778 ST66G08 P
33	77.2	30.4	469	24	A1726703	A1726703 BNGH1637
34	77.2	30.4	573	39	AW042658	AW042658 ST73A07 P
35	77.2	30.4	601	39	AW043146	AW043146 ST79H08 P
36	77.2	30.4	700	93	AW043255	AW043255 ST1B10 P
37	77.2	30.4	775	38	AW068560	AW068560 GA_Ea001
38	77.2	30.4	889	94	AW010325	AW010325 ST04G01 P
39	77.2	30.4	889	94	AW731435	AW731435 GA_Ea003
40	77.2	30.4	973	94	AW731454	AW731454 GA_Ea003
41	76.2	30.0	250	109	BE529534	BE529534 M74H09STM
42	76.2	30.0	384	37	AV551838	AV551838 AV551838
43	76.2	30.0	549	39	AW042889	AW042889 ST26F06 P
44	76.2	30.0	565	37	AV550681	AV550681 AV550681
45	76.2	30.0	568	37	AV549801	AV549801 AV549801

ALIGNMENTS

RESULT 1
 LOCUS AL369727 477 bp mRNA
 DEFINITION MBLA32H11F1 MBLA Medicago truncatula cDNA clone MBLA32H11 T3, mRNA
 ACCESSION AL369727
 VERSION AL369727.1 GI:9669480
 KEYWORDS EST.
 SOURCE barrel medic.

ORGANISM *Medicago truncatula*
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta
Magnoliophyta; eudicotyledons, core eudicots; Rosidae; eurosids I;
Fabales; Fabaceae; Papilionoideae; Medicago.
REFERENCE
1 (bases 1 to 477)
Journet, E. P., Crespeau, H., van Tuinen, D., Gouzy, J., Tallon, O.,
Niebel, A., Carreau, Y., Chataigner, O., Kahn, D., Glanina, Z., Pearson
V. and Gamas, P.
Medicago truncatula ESTs from nitrogen-starved roots
Unpublished (2000)
COMMENT
Title: JOURNAL
Contact: Genoscope
Genoscope - Centre National de Séquençage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Contact : Pascal Gamas and Etienne-Pascal Journet, Laboratoire de
Biologie Moléculaire des Relations Plantes-Microorganismes,
CNRS-INRA, BP 27 31326 Castanet-Tolosan Cedex, France (Email :
mc-gest@toulouse.inra.fr Website :
<http://sequence.toulouse.inra.fr/Mtruncatula.html>).

```

FEATURES
source
location/Qualifiers
1. 477
/organism="Medicago truncatula"
/cultivar="Jemalong"
/db_xref="taxon:3880"
/clone="MEBA32H11"
/clone_1fb="MCBA"
/tissue_type="root tips"
/dev_stage="harvested after 3 days of N-starvation"
/note="Vector: pBluescript PSK. Site_1: EcoRI; Site_2:
XhoI. Plants were grown in an aeroponic chamber for 14
days on nitrogen-rich medium followed by 3 days on N-free
medium. RNA was extracted from root tips (1-3 cm). cDNA
was prepared from polyA+ enriched RNA. The cDNA was
directionally ligated into Uni-ZapXR vector from
Stratagene and packaged using GigaPack Gold packaging
extracts. Plasmids containing cDNA inserts were
mass-excised from phage stocks using ExsacII helper phage
and propagated in SOLR cells. Clone ordering and
sequencing was performed by the Centre National de
Sequencage (Genoscope, Evry, France)."
BASE COUNT
140 a
85 c 95 g 157 t
ORIGIN

```

Query Match	50.2%	Score 127.6	DB 28	Length 477
Best Local Similarity	78.7%	Pred. No. 4e-29		
Matches 148	Conservative	0	Mismatches 40	Indels 0
				Gaps 0

[illegible]

RESULT	2
BE204104	
LOCUS	
DEFINITION	EST96780 KVO Medicago truncatula CDNA clone PKV0-14B21, mRNA
ACCESSION	BE204104
VERSION	BE204104.1
KEYWORDS	GI:8747387
EST.	

SOURCE	barrel medic.
ORGANISM	Medicago truncatula
REFERENCE	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Medicago.
AUTHORS	1 (bases 1 to 604) Vandenbosch, K., Endre, G., Hur, J., Moore, J., Beremand, P., Ellis, L., Town, C.D., Bowman, C.L., Craven, M.B., Hansen, T.S., Holt, I.E. and Fraser, C.M.
TITLE	ESTs from uninoculated seedling roots of Medicago truncatula
JOURNAL	Unpublished (1999)
COMMENT	Contact: Vandenbosch K Department of Biology Texas A&M University College Station, TX 77843-3258, USA Tel: 409 845 7707 Fax: 409 845 2891 Email: kate@mail.bio.tamu.edu Texas A&M University name:1263876e TIOR sequence name:MTGA0117K More information is available at http://chrysie.tamu.edu/medicago Seq primer: SKmod (CTA GAA CTA gta gta GAT CC). Location/availability
SEQUENCE	

```

FEATURES
source
    location/Qualifiers
    1..604
    /organism="Medicago truncatula"
    /cultivar="genotype A17"
    /db_xref="taxon:3880"
    /clone="pKV0-14B21"
    /clone_1fb="KV0"
    /tissue_type="Seedling roots"
    /dev_stage="Immediately prior to inoculation with
    Sinorhizobium meliloti (0 hour)"
    /lab_host="E.coli strain XL0R"
    /note="Vector: pBluescript SK -; Site:1: EcoRI; Site:2:
    XhoI; cDNA was prepared from polyA+ enriched RN. The cDNA
    was directionally ligated into the unizap XR vector from
    StrataGene and packaged using Gigapack III Gold packaging
    extracts. Plasmids containing cDNA inserts were excised
    from the recombinant lambda-Zap phage using Ex-assist
    helper phage and propagated in XL0R cells."

```

Query Match	50.2%	Score 127.6	DB 105	Length 604
Best Local Similarity	78.7%	Pred. No. 4.3e-29		
Matches 148	Conservative	0	Mismatches 40	Indels 0
				Gaps 0

OY	67	ctactgnccttcctcccnacacctgaaccctcaatcgcgtcgagatttcttcgaaga	126
Db	13		
OY	127	gattcgaagaatgta tggaaagcccgtcgtgctanccgaatcgaaanaagagtgaagaa	186
Db	73		
OY	187	aagcggblacctccaagcacacagcagaagaaaatvgctctggagatccngatgacnaagtc	246
Db	133		
OY	247	ttagaca 254	
Db	193	tttagaca 200	

RESULT	3
BE204201	628 bp mRNA
LOCUS	EST939687 KVO Medicago truncatula cDNA clone PKV0-14F14, mRNA
DEFINITION	sequence.
ACCESSION	BE204201
VERSION	BE204201.1 GI:8747485

KEYWORDS
SOURCE barrel medic.
ORGANISM Medicago truncatula
REFERENCE Eukaryota: Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Medicago.
1 (bases 1 to 628)
AUTHORS Vandenbosch, K., Endre, G., Hur, J., Moore, J., Beremand, P., Ellis, L., Town, C.D., Bowman, C.L., Craven, M.B., Hansen, T.S., Holt, I.E. and Fraser, C.M.
TITLE ESTs from uninoculated seedling roots of Medicago truncatula
JOURNAL Unpublished (1999)
COMMENT Contact: Vandenbosch K
Department of Biology
Texas A&M University
College Station, TX 77843-3258, USA
Tel: 409 845 7707
Fax: 409 845 2891
Email: kate@mail.bio.tamu.edu
Texas A&M University name: T263973e
TIGR sequence name: MTGAP317K
More information is available at:
http://chrysis.tamu.edu/medicago
Seq primer: SKmed (CTA gAA CTA gTg gAT CC).
FEATURES
source
1..628
Location/Qualifiers
/organism="Medicago truncatula"
/cultivar="genotype A17"
/db_xref="taxon:3880"
/clone="pKV0-14F14"
/clone_lib="KV0"
/tissue_type="Seedling roots"
/dev_stage="Immediately prior to inoculation with Sinorhizobium meliloti (0 hour)"
/lab_host="E.coli strain XL0LR"
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the Unizap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-assist helper phage and propagated in XL0LR cells."

BASE COUNT
188 a 111 c 134 g 195 t

ORIGIN
Query Match
Best Local Similarity 78.7%; Pred. No. 4.3e-29; Length 628;
Matches 148; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
50.2%; Score 127.6; DB 105;
Query 67 ctacttgccttcctcnaactcgtcacctcctcattcgtcgtgagatttcttcgaagaga 126
Db 12 CTCTTTCCTTCAATATACCTTTTCATCTCTCATTTTGTGCAAGTATTTTTCGAAGAGA 71
Qy 127 gattcgaagatgagatggaagagcgntggtgctcancgattggaagagtggaagaa 186
Db 72 CATTCCAAGATGGGTGGAAGAGTGGTGTATATCTGATTGGAAGAGTGAAGCAA 131
Qy 187 aagcgggtacctcgaacacagcagcaaaaatgctcgggagatccnatacnaagatc 246
Db 132 AAGCAGGAGTCTTATAGTATCTGCTGGAATAATGGCTGCTGATCTGATGACCAAGGTA 191
Qy 247 ttccagaca 254
Db 192 TTTCAGACA 199

RESULT 4
AM725203 948 bp mRNA EST 20-Apr-2000
LOCUS GA_Ea0015H01 Gossypium arboreum 7-10 dpa fiber library Gossypium
DEFINITION
arboresum cDNA clone GA_Ea0015H01, mRNA sequence.
ACCESSION AM725203

VERSION AM725203.1 GI:7622705
KEYWORDS EST.
SOURCE Gossypium arboreum.
ORGANISM Gossypium arboreum
REFERENCE Eukaryota: Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
1 (bases 1 to 948)
AUTHORS Leslie, A., Frisch, D., Yu, Y., Wood, T.C., Wing, R.A. and Wilkins, T.A.
TITLE An integrated analysis of the genetics, development, and evolution of the cotton fiber
JOURNAL Unpublished (2000)
COMMENT Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
High quality sequence stop: 948.
FEATURES
source
1..948
Location/Qualifiers
/organism="Gossypium arboreum"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ea0015H01"
/clone_lib="Gossypium arboreum 7-10 dpa fiber library"
/tissue_type="Fibers isolated from bolls harvested 7-10 dpa"
/lab_host="E. coli"
/note="Vector: pBK-CMV; Site 1: EcoRI; Site 2: XhoI"

BASE COUNT
228 a 234 c 209 g 265 t 2 others

ORIGIN
Query Match
Best Local Similarity 47.0%; Score 119.4; DB 94; Length 948;
Matches 141; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
77.0%; Pred. No. 1.7e-26;
Query 71 ttgpttcttcctcnaactcgtcacctcctcattcgtcgtgagatttcttcgaagagatt 130
Db 257 TTTTCTTCACTTCTACCTCTTCATTTGCTCTCTAGATTTCTTCCAGAACGCTT 316
Qy 131 cgaagatgatggaagagcgntggtgctcancgattggaagagtggaagaaagc 190
Db 317 CAAGATGTGGCGGAGAGACCGCATGGGTCTCTCAATTTGGAAGAGTGAACGAAAGC 376
Qy 191 gggtacctcgaacacagcaggaagaatggtcgtggagatccngatgaacnaagtcctca 250
Db 377 CGGCACCTTTAAGCACACCGCGCAAGTGTCTGGCATCTCATGATTAAGGTATTCGA 436
Qy 251 gac 253
Db 437 GAC 439

RESULT 5
BE462089 463 bp mRNA EST 27-JUL-2000
LOCUS BE462089 tomato flower buds 0-3 mm, Cornell University
DEFINITION
Lycopersicon esculentum cDNA clone cTOA11113, mRNA sequence.
ACCESSION BE462089
VERSION BE462089.1 GI:9507858
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
REFERENCE Eukaryota: Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 463)
AUTHORS van der Hoeven, R.S., Bezzerides, J.L., Matern, A.L., Holt, I.E., Liang, F., Hansen, T., Craven, M.B., Bowman, C.L., Ronning, C.M., Nierman, W.,

TITLE Fraser, C.M., Matlin, G.B., Giovannoni, J.J. and Tanksley, S.D.
 Generation of ESTs from tomato flower tissue, 0-3 mm buds
 JOURNAL Unpublished (1999)
 COMMENT Contact: David Frisch

TITLE
Graser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,S.D
JOURNAL
Generation of ESTs from tomato flower tissue, 0-3 mm buds
COMMENT
Unpublished (1999)
Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU
5 prime sequence.

FEATURES	Location/Qualifiers
source	1. .463

/organism="Lycopersicon esculentum"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="cMOA11113"
 /clone_1ib="tomato flower buds 0-3 mm, Cornell University"
 /tissue_type="flower"
 /dev_stage="0-3mm buds"
 /note="vector: pBluescript SK(-); site_1: EcoRI, site_2:
 XhoI; supplier: Tanksley; Flower buds and flowers were
 taken from greenhouse plants (4-8 wk old, TA496). They
 were immediately frozen in liquid nitrogen and then
 size-separated while remaining frozen."

Query Match	43.5%	Score	110.4	DB	108	Length	463
Best Local Similarity	71.2%	Pred. No.	8.6e-24				
Matches	141	Conservative	0	Mismatches	57	Indels	0
						Gaps	0

Oy 56 atgcgctcaagctacttgnctttcccttcnaacgtcgacatccttaattcgctgaagattc 115
||| ||| | | | | | | | | |
Db 26 atcatTTTCTACTAGATTGTGTAAGTTTCATTAATTACATCTTCATTAATCGAGACTT 85

Oy 116 ctctcgaagagagatctcgaaagtggatgataagaagccgnltgggtgcctnccgatltgaaaag 175
||| ||||| ||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||||
Db 86 ctttTGACGACAGCAATTTTCATCATGTATGGCAGACAGATGGCTCGACTCTGCATGGAAAG 145

Qy 176 gaaTgaaagaaaaagcgggtaccttcaagcacacagcagcaaaatgltcgygatccnga 235

Db 146 GAGTGAAGGGAACGACGGCTCTTTAAGCATACAGCTGGGAATGGGCTGGTGATCTTGA 205

Oy	236	tgacnaaggtcttcagac	253
Dh	206	TGATAAGGTATTTCAGAC	223

RESULT 6
BE921143

LOCUS	02-001-2000
DEFINITION	EST42412 potato leaves and petioles Solanum tuberosum cDNA clone
ACCESSION	BE921143

VERSION	01.1044/219
KEYWORDS	EST.
SOURCE	potato.
ORGANISM	Solanum tuberosum

REFERENCE
1 (bases 1 to 468)
I; Solanales; Solanaceae; Solanum.

TITLE	AUTHORS
Generation of ESTs from potato leaves and petioles	van der Hoeven, R.S., Bezzerides, J., Holt, L.E., Liang, F., Cho, J.
	Uterback, T., Hansen, C.L., Doan, B., Bougrit, O., Buell, C.R., Roman, C.M., Fry, W.E., Tanksley, S.D. and Baker, B.

COMMENT
The Institute for Genomic Research
For clone request: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com.
contact: carny kornblum

FEATURES	Location/Qualifiers
source	1. .468

```

FEATURES
SOURCE

Location/Qualifiers
1. 468
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:413"
/clone="CSFRLIN24"
/clone_1lb="potato leaves and petioles"
/tissue_type="leaflets and petioles"
/dev_stage="8 weeks old plants"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Tissue was supplied by Dr. Fry (Cornell University)
leaflets and petioles were isolated from 8 week old
greenhouse grown plants. The plants were watered and
fertilized freely. The tissue was immediately frozen in
liquid nitrogen."

```

BASE COUNT	141 a	80 c	104 g	143 t
ORIGIN				

Query MatchScore	43.58;	Score	110.4;	DB	137;	Length	468;
Best Local Similarity	71.28;	Pred. No.	8.6e-24;				
Matches	141;	Conservative	0;	Mismatches	57;	Indels	0;
						Gaps	0;

Qy 56 atgcgtcgaagctacttgncttccttccttcnaaccltgcaactcttcattcgccgagatltt 11
||| ||| | | | | | | | | | |
Db 26 ATCAATTCTACTAGTTTGTTCCTGTTTCAATTAACACATCTTCATTAATCGAGATCCT 85

Oy 116 cttcgaagagaagatctcgaaagtgaatgaaagaccnctggctgcclanccgatctgaaaag 177
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 86 ctttgaagacatttgatgatgttcgacagacgatgcctgaagctgactcggaag 147

Oy 176 gacgcgaagaaagcgcggtacccttcacaacagcaggaataatgtctcggtatccnga 23
| | | | | | | | | | | | | | | | | | | | |
Db 146 GACGCAAGGAAAGCAGCGCTCCCTTAACATACAGCTGGGAATGCCTGCATCATCCGTA 20

Qy	236	tgacnaagctctcagac	253
Db	206	TGATAAGCTATTCAgac	223

RESULT 7
AM041721

DEFINITION ES228454585 tomato mixed elicitor, BTI Lycopersicon esculentum c
clone CLETT14M2, mRNA sequence.

FUNCTION	EN041/22.1	01.0004/2
KEYWORDS	EST,	
SOURCE	tomato.	
ORGANISM	Lycopersicon	esculentum

REFERENCE
1 (bases 1 to 263)

ASCHECO, M., HE, A., LYMAN, J., HOL, L., LAUNG, F., UPTON, J.,
 RONNING, C.M., CRAVEN, M.B., FUJII, C.Y., BOWMAN, C.L.,
 NIEMAN, W., FRASER, C.M., VENTER, J.C., MARTIN, G.B., TANKSLEY, S.D.,
 and GLOVER, J.

TITLE	Generation of ESis from Committed Tissue
JOURNAL	Unpublished (1999)
COMMENT	Contact: David Frisch Clemson University Genomics Institute

Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293

```

FEATURES
source      1 263
            location/qualifiers
            5 prime sequence.
            email: dirisch@CLEMSON.EDU

```

```
/organism="Lycopersicon esculentum"  
/cultivar="Rio Grande Ptor"
```

```

/db_xref="taxon:4081"
/clone="cLET14M2"
/clone_lib="tomato mixed elicitor, BRT"
/tissue_type="leaf"
/dev_stage="4-6 week old plants"
/lab_host="XLI-Blue MRP"
/notes="Vector: pBluescript SK(-); Site.1: EcoRI; Site.2: XhoI; cLET - inoculated with a variety of disease response elicitors. Plants exposed to 2,6 dichloroisonicotinic acid, BTH, jasmonic acid, ethylene, fenchone, EIX, okadaic acid, or systemin prior to tissue harvest. EcoRI site was destroyed during cloning."
BASE COUNT      81 a      50 c      56 g      76 t
ORIGIN

Query Match      42.8%; Score 108.8; DB 39; Length 263;
Best Local Similarity 70.7%; Pred. No. 2.3e-23;
Matches 140; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

56 atcgcctcaagctactgcttcctccctcnaactctgcactcttcactgcgtgatctt 115
   ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
11 AWCATTTCTACTAAATTTGGTTACTATTGCAATTTATACATGTCATTCATTCGAGATCT 70

0Y 116 cttcgaaagaagactcgaaatgagtgtgaagaagccgnitgggtgcanccgatgtgaaag 175
   ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 71 CTTTGAGAGACATTGTGATGATGATTTGGACGACAGATGGCGAAGACTGACTGAGAAAAG 130

0Y 176 gagtgaagaaagcgggttaccttcaagcacacacagaaatggtctgggagtcnga 235
   ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 131 GAGTGAAGGAAGAAGAGCGCTCTTTAAGCATATACACTGGCAATATGGCGTGTGATCTCTGA 190

0Y 236 tgaacnaaggtcttcagac 253
   ||| ||| | | | | | | | |
Db 191 TGATTAAGGTATTTCATYAC 208

RESULT 8
BE449253 608 bp mRNA EST 26-JUL-2000
DEFINITION EST356108 L. hirsutum trichome, Cornell University Lycopersicon
hirsutum cDNA clone cLHT31A6, mRNA sequence.
ACCESSION BE449253
VERSION BE449253.1 GI:9454852
KEYWORDS EST.
SOURCE Lycopersicon hirsutum.
ORGANISM Lycopersicon hirsutum.
REFERENCE 1 (bases 1 to 608)
AUTHORS van der Hoeven,R.S., Bezzeredes,J.L., Matern,A.L., Holt,I.E., Liang
,F., Hansen,T., Craven,M.B., Bowman,C.L., Roming,C.M., Nierman,W.,
Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
TITLE Generation of ESTs from wild tomato (Lycopersicon hirsutum)
trichomes
JOURNAL Unpublished (2000)
COMMENT Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29654, USA
Tel: 864 656 4366
Fax: 864 656 4393
Email: dfrisch@clemson.edu
5 prime sequence.
FEATURES
Source
location/Qualifiers
1..608
/organism="Lycopersicon hirsutum"
/db_xref="taxon:62890"
/clone="cLHT31A6"
/clone_lib="L. hirsutum trichome, Cornell University"
/tissue_type="trichome"
/dev_stage="mixed stages"

```

BASE COUNT	188 a	113 c	135 g	172 t
ORIGIN				
Query Match	42.8%	Score 108.8;	DB 108;	Length 608;
Best Local Similarity	70.7%;	Pred. No. 2.9e-23;		
Matches 140:	Conservative	0;	Mismatches 58;	Indels 0; Gaps 0;
OY	56 atcgctcaagctaacttgcatttcctccctcnaactlgtcacctlgtcatcgtcgatgattt	115		
Db	26 ATCATTTCTACTACAGTTTTGTTTACTCTTTTCATTAATAATACATCCTTCATTTCGAATCTT	85		
OY	116 ctltgaagaagagatctcgagaagtgaagtgaagaagccgnhtgggtgtcancncgatltgaaaaa	175		
Db	86 CTTTAAGAAGACAATTTGTATGATGATGGCAGAGCGAGATGGTGTAAGCTGTGATGGAAAG	145		
OY	176 gagtgaagaaagagggttaccttccaagcacacaacagcagaaaatgtctggtgatacunga	235		
Db	146 GAGTGAAGGAAAGCAGAGCGCTCTTAAGCATACACTGGGAATATGGCTGTGATCCCTGA	205		
OY	236 tgacnaagglctcagac	253		
Db	206 TGATTAAGCTATTTCAGAC	223		
RESULT 9				
LOCUS	AW934135	529 bp	mRNA	EST
DEFINITION	EST359978 tomato fruit mature green, TAMU Lycopersicon esculentum			
ACCESSION	CDNA Clone CLEF57P10 5', mRNA sequence.			
VERSION	AW934135			
KEYWORDS	AW934135.1 GI:8109536			
SOURCE	EST.			
ORGANISM	tomato.			
TITLE	Lycopersicon esculentum			
JOURNAL	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon. 1 (bases 1 to 529)			
AUTHORS	Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Jiang,F., Upson,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Romling,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J. Generation of ESTs from tomato fruit tissue Unpublished (1999) Contact: David Frisch Clemson University Genomics Institute 100 Jordan Hall, Clemson, SC 29634, USA Tel: 864 656 4366 Fax: 864 656 4293 Email: dfrisch@CLEMSON.EDU 5 prime sequence.			
COMMENT	Location/Qualifiers			
FEATURES	1..529			
SOURCE	/organism="Lycopersicon esculentum" /cultivar="TA496" /db_xref="taxon:4081" /clone="CLEF57P10" /clone_lib="tomato fruit mature green, TAMU" /issue_type="fruit pericarp" /dev_stage="mature green (3-5 days pre-ripening)" /lab_host="SOLR" /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; clef - Fruit were tagged at the im stage and harvested 3-5 days prior to ripening. Fruit were cut in half to verify the seeds were indeed immature and the seeds and locules were discarded prior to freezing the pericarp"			

ORIGIN	BASE COUNT	165 a	96 c	119 g	149 t
Query Match	42.4%	Score 107.8;	DB 97;	Length 529;	
Best Local Similarity	75.6%;	Pred. No. 5,7e-23;			
Matches 130;	Conservative	0;	Mismatches 42;	Indels 0;	Gaps 0
Oy	82	ttcnaacctgcacaccttcattcgcgttgatgttcttcgaaagagatcgcgaagatcgcgaatgat	141		
Db	52	ttttcaattatattacattcttcattatctgaatcttcttttgaaacacaaatttgatgatgatt	111		
Oy	142	ggaagaagccgntgggtgcctcancagacttcgaaagaagatgaagaagaagctggcttacc	201		
Db	112	ggcagacgacgagatgcgtgaagctgactgcaaaagacgtaaacgaaagcagacgctctctt	171		
Oy	202	agcacacagcaggaataatgctctgggtggtcncgatgcgaagctgtcttcagac	253		
Db	172	agcattacagctgggaattggcgtgctgatctgatgatgtaaggtatttcagac	223		
BLAST 10					
BE420773					
LOCUS	BE420773	807 bp	mRNA	EST	24-JUL-2000
DEFINITION	HM002.D06 ITEC HMM Barley Leaf Library Hordeum vulgare cDNA clone				
ACCESSION	HM002.D06				
VERSION	BE420773				
KEYWORDS	BE420773.1	GI:9418616			
SOURCE	EST.				
ORGANISM	Hordeum vulgare				
REFERENCE	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta				
AUTHORS	Magnoliophyta; Liliopsida; Poales; Poaceae; Hordeum.				
	1 (bases 1 to 807)				
	Anderson,O.A., Appels,R., Bailey,P., Blake,T., Close,T., Cloutier				
	,S., Dubcovsky,D., Feuillet,C., Gale,M., Graner,A., Gustafson,P.,				
	Herrmann,R.G., Holton,T., Jacquemin,J.M., Jia,J., Joudrier,P.,				
	Langridge,P., Lazo,G.R., Lin,J.J., McGuire,P., Ogihara,Y.,				
	Pechion,L.N., Qualset,C., Schuch,W., Selvaraj,C., Shariflou,M.,				
	Sorrells,M., Warburton,M. and Wenzel,G.				
	International Triticeae EST Cooperative (ITEC): Production of				
	Expressed Sequence Tags for Species of the Triticeae				
	Unpublished (2000)				
JOURNAL	Contact: Herrmann RG				
COMMENT	Botanisches Institut der LMU				
	Menzinger str. 67, D-80638 Muenchen GERMANY				
	Fax: 49 30 171683				
	Email: herrmann@botanik.biologie.uni-muenchen.de				
	International Triticeae EST Cooperative (ITEC)				
	http://wheat.pw.usda.gov/genome.				
FEATURES	location/Qualifiers				
SOURCE	1. 807				
	/organism="Hordeum vulgare"				
	/cultivar="Barke"				
	/db_xref="taxon:4513"				
	/clone="HM002.D06"				
	/clone_lib="ITEC HMM Barley Leaf Library"				
	/tissue_type="leaf"				
	/dev_stage="14 day old"				
	/note="Vector: pBluescriptSK(-); 850 bp average insert				
	size."				
BASE COUNT	210 a	206 c	215 g	170 t	6 others
ORIGIN					
Query Match	42.4%	Score 107.8;	DB 108;	Length 807;	
Best Local Similarity	80.0%;	Pred. No. 6.4e-23;			
Matches 124;	Conservative	0;	Mismatches 31;	Indels 0;	Gaps 0
Oy	100	cattgcgcgaagcttcttccttcgaagagatcgaagaatgcgaagaagcagcgttgctc	159		
Db	248	cctccggggaggtatcttccagagacgcttcgaaagatggttgagacacgttgagta	307		

QY	160	tancgatttgcaaaagagtgtaaggaagaaagcggttaccctcaagacacagaggaataat	219
Db	308	AATCGGATTGGAAAAAGGAGCGAAGGAAAGCCGGTACATTCAAGCACACAGCGAGGAAAT	367
QY	220	ggtctggagatccngatgacnaaggctctcgaca	254
Db	368	ATTCTGGGATCTCGATGACAAAGGCATTCAACACA	402

RESULT	11		
BE193174			
LOCUS			
DEFINITION	BE193174 714 bp mRNA EST 25-JUL-2000		
ACCESSION	HYSMEN0080A07F Hordeum vulgare 5-45 DAP spike EST library		
VERSION	HYCDNA0009 (5 to 45 DAP) Hordeum vulgare cDNA clone HYSMEN0080A07F		
KEYWORDS	mRNA sequence.		
SOURCE	BE193174		
ORGANISM	BE193174.1 GI:8705353		
REFERENCE	EST.		
AUTHORS	barley.		
	Hordeum vulgare		
	Eukaryote; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;		
	Magnoliophyta; Liliopsida; Poales; Poaceae; Hordeum.		
	1 (bases 1 to 714)		
	Wing, R., Close, T.J., Klein, H., Wise, R., Begum, D., Frisch, D., Y.		
	'T., Saski, C., Schwartzbeck, J., Simmons, J., Choi, D.W., Main, D. and		
	Wood, T.		
	Development of a genetically and physically anchored EST resource		
	for barley genomics		
	Unpublished (2000)		
JOURNAL	Contact: Wing RA		
COMMENT	Clemson University Genomics Institute		
	Clemson University		
	100 Jordan Hall, Clemson, SC 29634, USA		
	Tel: 864 656 7288		
	Fax: 864 656 4293		
	Email: rwing@clemson.edu		
	Seq primer: AATTACCCCTACTTAAAGG		
	High quality sequence start: 187		
	High quality sequence stop: 454.		
FEATURES	Location/Qualifiers		
SOURCE	1..714		
	/organism="Hordeum vulgare"		
	/cultivar="Morex"		
	/db_xref="taxon:4513"		
	/clone="HYSMEN0080A07F"		
	/clone_lib="Hordeum vulgare 5-45 DAP spike EST library		
	HYCDNA0009 (5 to 45 DAP)"		
	/rname_type="5-45 DAP Spike"		
	/lab_host="SOLR"		
	/note="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI"		
BASE COUNT	192 a 188 c 183 g 150 t 1 others		
ORIGIN			
Query Match	41.2% Score 104.6; DB 105; Length 714;		
Best Local Similarity	78.7%; Pred. No. 6.1e-22;		
Matches 122; Conservative 0; Mismatches 33; Indels 0; Gaps 0;			
QY	100	cattcgctgaagattctctcgaagagagatcgaaagatgagatggaagacgntgggtgc	159
Db	123	CCCTCGGGGAGTCTCTTCGAAAGCGGTTTCGAAGATGTTGGGAGACACTTGGGTCA	182
QY	160	tancgatttgcaaaagagtgtaaggaagaaagcggttaccctcaagacacagaggaataat	219
Db	183	AATCGGATTGGAAAAAGGAGCGAAGGAAAGCCGGTACATTCAACACACAGCGAGGAAAT	242
QY	220	ggtctggagatccngatgacnaaggctctcgaca	254
Db	243	ATTCTGGGATCTCGATGACAAAGGCATTCCACACA	277

RESULT	12
--------	----

BE598903
LOCUS BE598903 547 bp mRNA EST 18-AUG-2000
DEFINITION P11_83_A08.b1_A002 Pathogen induced 1 (P11) Sorghum bicolor cDNA, mRNA sequence.
ACCESSION BE598903
VERSION BE598903.1 GI:9853962
KEYWORDS EST.
SOURCE sorghum.
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Sorghum.
1 (bases 1 to 547)
Cordonnier-Pratt,M.-M., Gingle,A., Dean,R., Sudman,M. and Pratt,L.H.
REFERENCE An EST database from Sorghum: pathogen-induced plants
AUTHORS Unpublished (2000)
JOURNAL Contact: Cordonnier-Pratt MM
COMMENT Department of Botany
The University of Georgia
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 542 1805
Email: empratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.
Seq primer: JEN REV
High quality sequence stop: 478
POLYA-No.
FEATURES
source Location/Qualifiers
1..547
/organism="Sorghum bicolor"
/db_xref="taxon:4558"
/clone_id="Pathogen induced 1 (P11)"
/note="Organ: Anthracnose induced 48 hr after inoculation: two-week-old sorghum plants 48 hr after inoculation: Vector: pBluescript II from Lambda Zap II; Site1: XhoI; Site2: EcoRI; Two-week-old sorghum plants (BTX 623 cultivar) were infected with pathogen (isolate FRM421 of Colletotrichum graminicola, which is a sorghum isolate). RNA was prepared from infected leaves harvested from 45 seedlings 48 hours after inoculation. Note: Young seedlings (2 weeks old) exhibit juvenile resistant reaction, which is an incompatible interaction. As they grow older (4 weeks or older), plants resume susceptibility to anthracnose disease. The library was made from poly-A RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision. WARNING: While most or all ESTs are expected to derive from the host plant, no effort was made to eliminate ESTs deriving from the pathogen."

BASE COUNT 136 a 142 c 142 g 127 t
ORIGIN
Query Match 40.7%; Score 103.4; DB 110; Length 547;
Best Local Similarity 75.8%; Pred. No. 1.3e-21;
Matches 125; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

OY 90 ctgaccccttcattcgcctgagatttcttcgaagaagatcgaagaatgagatgaagc 149
|||||
DB 116 CTGCTCCCTCGCTTCGGGTGAATGCTCTTCAAGAGCGCTTCGAAGATGTTGGAGAGT 175
OY 150 cgnatggtgtctancgatctggaagaaagatgaagaaagcgggtactctcaagaca 209
|||||
DB 176 CGGTGGGTGAAGATCGAATGGAAGAAAGGATGAAGGAAAGCTGGAAGTTCAACACACA 235
OY 210 gcagagaaaatggtctgggagatcngatgaacnaaggtcttcagaca 254
|||||
DB 236 GCGGGGAGATCTCTGCAGATCTGATGACCAAGGAATACAAACA 280

RESULT 13

AT002063
LOCUS AT002063 290 bp mRNA EST 10-AUG-1999
DEFINITION AT002063 Flower bud cDNA Brassica rapa subsp. pekinensis cDNA clone RF0337, mRNA sequence.
ACCESSION AT002063
VERSION AT002063.1 GI:5724847
KEYWORDS EST.
SOURCE Chinese cabbage.
ORGANISM Brassica rapa subsp. pekinensis
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 290)
Ryu,S.W., Lim,C.O. and Cho,M.J.
REFERENCE Brassica EST (Flower bud cDNA)
AUTHORS Unpublished (1999)
JOURNAL Contact: Cho MJ
COMMENT Department of Biochemistry
Gyeongsang National University
#900 Gajwa-dong, Chinju, Gyeongnam 660-701, Korea
Tel: 82-591-751-5957
Fax: 82-591-759-9363
Email: mjcho@ngae.gsnu.ac.kr
Submitted through BRIC(Biological Research Information Center) of Korea
URL: <http://bric.postech.ac.kr/>.
FEATURES
source Location/Qualifiers
1..290
/organism="Brassica rapa subsp. pekinensis"
/db_xref="taxon:51351"
/clone_RF0337"
/clone_id="Flower bud cDNA"
/lab_host="MC1061"
/note="Vector: pBluescript KS(+); Site1: EcoRI; Site2: XhoI; average insert size:700 bp; initial pfu:5.400 x 10⁹"
BASE COUNT 85 a 70 c 60 g 75 t
ORIGIN
Query Match 40.5%; Score 102.8; DB 28; Length 290;
Best Local Similarity 62.9%; Pred. NO. 1.7e-21;
Matches 149; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

OY 17 ctaaggaatgnaatgnaatgagcaannccgaacacagtcagctcgaagtactgct 76
|||||
DB 4 CTGACAAAATGATTAACCAACAAACAGCTCAAGCTTCCACCTTCTCTCTC 63
OY 77 ttccctcnaactcgcacactccttcattcgtgagatttcttcgaagaagatcgaaga 136
|||||
DB 64 TCTTCTCACTTAACCTCCCTGGCTTCTCTGAGATCTTCTTGAAGACATTTGGAAGG 123
OY 137 tggatggaagagcgcgttggtggtctancgcatttggaagaagatgaagaagcggtac 196
|||||
DB 124 TGGATGGAAGAGCAGGTGGCTTATCTGATTGGAAGAAACGAAGCAAGCAAGCTGTAC 183
OY 197 ctcaagacacacgaaggaagaaatggtctgggagatcngatgaacnaaggtcttcagac 253
|||||
DB 184 CTTTAAGCACAACCGCGGCAAAATGGCCCGCATGCCGATCCGACATTAAGTATCCAGAC 240

RESULT 14
AT001761
LOCUS AT001761 411 bp mRNA EST 14-JUN-1999
DEFINITION AT001761 Flower bud cDNA Brassica rapa subsp. pekinensis cDNA clone RF0156, mRNA sequence.
ACCESSION AT001761
VERSION AT001761.1 GI:5056232
KEYWORDS EST.
SOURCE Chinese cabbage.
ORGANISM Brassica rapa subsp. pekinensis
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;

REFERENCE 1 (bases 1 to 411)
 AUTHORS Ryu,S.W., Lim,C.O. and Cho,M.J.
 TITLE Brassica EST (Flower bud cDNA)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Cho MJ
 Department of Biochemistry
 Gyeongsang National University
 #900 Gajwa-dong, Chinju, Gyeongsang 660-701, Korea
 Tel: 82-591-751-5957
 Fax: 82-591-759-9363
 Email: mjcho@ngae.gsnu.ac.kr
 Submitted through BRIC(Biological Research Information Center) of Korea
 URL: http://bric.postech.ac.kr/.

FEATURES
 source Location/Qualifiers
 1..411
 /organism="Brassica rapa subsp. pekinensis"
 /db_xref="taxon:51351"
 /clone="RF0156"
 /clone_lib="Flower bud cDNA"
 /lab_host="MC1061"
 /note="vector: phagescript KS(+); Site_1: EcoRI; Site_2: XhoI; average insert size:700 bp; Initial pfu:5,400 X 10⁹"

BASE COUNT 121 a 104 c 82 g 104 t
 ORIGIN

Query Match 40.5%; Score 102.8; DB 28; Length 411;
 Best Local Similarity 62.9%; Pred. No. 1.9e-21;
 Matches 149; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 17 cttgaagaaatgnaatgccaanncgaacactgtatcgtctcaagctactgntc 76
 DB 4 ctgagaaaataatgagattttccgaagacagacgcttttgcgccttcgac 76
 QY 77 ttctctcnaactctggccttctctctctctctctctctctctctctctct 136
 DB 64 tcttctcactcttcttcccttgccttctctctctctctctctctctctct 123
 QY 137 tggatggaaagagccgntgggtctcancgaatggaaaagagtgaaagagcggtac 196
 DB 124 tggatggaaagagcagctggtcttctctctctctctctctctctctctct 183
 QY 197 cttcaagccacagcagcgaatggtctctgggagtcngatgcnaaggtcttcagac 253
 DB 184 ctttaagcaccgacggcggaatggcccgcatcccgacaaatgaatgataccagac 240

RESULT 15
 AM624478
 LOCUS 564 bp mRNA EST 28-MAR-2000
 DEFINITION EST32423 tomato flower buds 3-8 mm, Cornell University
 Lycopersicon esculentum cDNA clone cTOB15N2 5', mRNA sequence.
 ACCESSION AM624478
 VERSION AM624478.1 GI:7337505
 KEYWORDS EST.
 SOURCE tomato.
 ORGANISM Lycopersicon esculentum
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids
 I; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE 1 (bases 1 to 564)
 AUTHORS van der Hoeven,R.S., Bezzerides,J.L., Matern,A.L., Holt,I.E., Liang
 F., Hansen,T., Craven,M.B., Bowman,C.L., Ronning,C.M., Nierman,W.,
 Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
 TITLE Generation of ESTs from tomato flower tissue, 3-8 mm buds
 JOURNAL Unpublished (1999)
 COMMENT Contact: David Frisch
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 4366
 Fax: 864 656 4293
 Email: dfisch@CLEMSON.EDU
 5 prime sequence.

FEATURES
 source Location/Qualifiers
 1..564
 /organism="Lycopersicon esculentum"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="cTOB15N2"
 /clone_lib="tomato flower buds 3-8 mm, Cornell University"
 /issue_type="flower"
 /dev_stage="3-8mm buds"
 /note="vector: phagescript SK(-); Site_1: EcoRI; Site_2: XhoI; supplier: Tanksley; Flower buds and flowers were taken from greenhouse plants (4-8 wks old, TA496). They were immediately frozen in liquid nitrogen and then size-separated while remaining frozen."

BASE COUNT 191 a 100 c 128 g 145 t
 ORIGIN

Query Match 39.2%; Score 99.6; DB 93; Length 564;
 Best Local Similarity 74.1%; Pred. No. 2e-20;
 Matches 123; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 89 tctgcactctcatctgcctgaagattctctcgaagagagatcgaagatgagtgaaag 148
 DB 6 tctgacttcttctcagatgtagacaccttcttgaagacgattgtagatgtagtggcgaa 65
 QY 149 ccgntgggtctcancgaltggaagagtgaaagagaaagcggtacactcagacac 208
 DB 66 caaatgggtgaactctcagtgaggaagaaagcactgaagcgaaggaatttaaccacac 125
 QY 209 agcaggaaaaatggtctcgggagtcngatgcnaaggtcttcagaca 254
 DB 126 agcgggcaactggcctggtgctgctgacagagtgatgacagagcttccagaca 171

Search completed: March 18, 2001, 07:37:32
 Job time: 1289 sec

This Page Blank (intd)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 18, 2001, 07:22:38 ; Search time 53.84 Seconds

(Without alignments)
760.303 Million cell updates/sec

Title: US-09-199-129A-1

Perfect score: 254
Sequence: 1 ggaatcgaagctaccta.....atgaacaaagcttcagaca 254

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 280836 seqs, 80580151 residues

1 number of hits satisfying chosen parameters: 561672

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgn2_6/prodata/2/lna/5A_COMB.seq:*
2: /cgn2_6/prodata/2/lna/5B_COMB.seq:*
3: /cgn2_6/prodata/2/lna/6_COMB.seq:*
4: /cgn2_6/prodata/2/lna/PCrus_COMB.seq:*
5: /cgn2_6/prodata/2/lna/Backfile1.seq:*

Prd. No. is the number of results predicted
score greater than or equal to the specified
and is derived by analysis of the total score
SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	33.6	13.2	7218	1	US-08-232-463-14
2	30.8	12.1	2518	3	US-09-433-699-3
3	29.2	11.5	3046	1	US-08-726-725-1
4	28.4	11.2	500	2	US-08-967-101-47
5	28.4	11.2	500	2	US-08-592-541-47
6	28.4	11.2	500	3	US-09-124-698-47
7	28	11.0	3303	1	US-08-081-610-3
8	28	11.0	6395	2	US-08-687-559-2
9	28	11.0	6395	2	US-09-259-741-1
10	28	11.0	6395	3	US-09-037-751-1
11	28	11.0	6425	3	US-09-259-741-3
12	28	11.0	6425	3	US-09-037-751-3
13	28	11.0	6439	3	US-09-259-741-2
14	28	11.0	6439	3	US-09-037-751-2
15	28	11.0	6446	3	US-09-259-741-5
16	28	11.0	6446	3	US-09-037-751-5
17	28	11.0	6475	3	US-09-259-741-4
18	28	11.0	6475	3	US-09-037-751-4
19	27.6	10.9	8937	2	US-08-449-933-3
20	27.6	10.9	8937	2	US-08-449-933-1
21	27.6	10.9	10706	1	US-08-411-389-1
22	27.4	10.8	43676	3	US-09-356-952-12
23	26.8	10.6	1211	3	US-08-956-182-23
24	26.8	10.6	1248	3	US-08-956-182-39
25	26.8	10.6	1366	3	US-08-956-182-27
26	26.8	10.6	1307	3	US-08-956-182-40
27	26.8	10.6	7218	1	US-08-232-463-14
28	26.8	10.6	87350	3	US-08-781-891-79

C 29	26.6	10.5	4002	1	US-08-331-488A-1	Sequence 1, App1
C 30	26.4	10.4	780	1	US-08-325-553-27	Sequence 27, App1
C 31	26.4	10.4	780	2	US-08-394-152A-27	Sequence 27, App1
C 32	26.4	10.4	2939	2	US-08-560-398-11	Sequence 11, App1
C 33	26.2	10.3	340	3	US-08-441-971-21	Sequence 21, App1
C 34	26.2	10.3	340	3	US-08-441-971-22	Sequence 22, App1
C 35	26.2	10.3	975	3	US-09-015-754-1	Sequence 1, App1
C 36	26.2	10.3	1254	3	US-09-015-754-3	Sequence 3, App1
C 37	26.2	10.3	2721	3	US-08-921-195A-1	Sequence 1, App1
C 38	26	10.2	1157	3	US-08-215-083-1	Sequence 1, App1
C 39	26	10.2	1157	3	US-08-900-026-1	Sequence 1, App1
C 40	26	10.2	6605	1	US-08-769-309A-4	Sequence 4, App1
C 41	26	10.2	6605	3	US-08-994-570-4	Sequence 4, App1
C 42	25.8	10.2	846	3	US-09-154-874-1	Sequence 1, App1
C 43	25.8	10.2	4739	3	US-08-685-871-1	Sequence 1, App1
C 44	25.6	10.1	1395	3	US-08-467-023-140	Sequence 140, App1
C 45	25.6	10.1	1410	3	US-08-467-023-139	Sequence 139, App1

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232.463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935.313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29 768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZ9pt-F15
US-08-232-463-14

Query Match 13.2%; Score 33.6; DB 1; Length 7218;

```

Best Local Similarity 4.5%; Pred No. 0.032; Indels 0; Gaps 0
Matches 6; Conservative 85; Mismatches 43; Indels 0; Gaps 0

OY 111 attctctcgaagaagattcgaagtgaatgaaagccgnttgggtctanccgattgg 170
    |||| | : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1443 ATTGTGACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1384

OY 171 aaaaggaatgaaagaaagcggtgtccttcaagcacaagcagaagaatagtctcgggat 230
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1383 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1324

OY 231 ccngatgacnaag 244
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1323 RRRRRRRRRRRRRR 1310

RESULT 2
US-09-433-699-3
Sequence 3, Application US/09433699B
Patent No. 6165786
GENERAL INFORMATION:
APPLICANT: C. Frank Bennett
APPLICANT: Lex M. Cowsett
TITLE OF INVENTION: ANTISENSE MODULATION OF NUCLEOLIN EXPRESSION
FILE REFERENCE: RFS-0109
CURRENT APPLICATION NUMBER: US/09/433.699B
CURRENT FILING DATE: 1999-11-03
NUMBER OF SEQ ID NOS: 89
SEQ ID NO 3
LENGTH: 2518
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (112)..(2235)
US-09-433-699-3

Query Match 12.1%; Score 30.8; DB 3; Length 2518;
Best Local Similarity 51.6%; Pred. No. 0.19;
Matches 65; Conservative 0; Mismatches 61; Indels 0; Gaps 0.

OY 52 agtcacgtctcaagactactcgtgaccttctccttcnaactctgcaccttcattcgctgaga 111
    || | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1271 agaagagcggagatcgcggacaacttcttggctaaatactccttacaagaagtcactcag 1330

OY 112 ttcttcggaagagatctgaagatgtagaagaagccgmgtgtctancgattga 171
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
    1331 atgaattgaagaagatgtttgaaagatcgtcggagatcagattagtcagcaagatggga 1390

OY 172 aaagga 177
    ||| |
Db 1391 aaagta 1396

RESULT 3
US-08-726-725-1
Sequence 1, Application US/08726725
Patent No. 5773290
GENERAL INFORMATION:
APPLICANT: Gould, Michael N.
APPLICANT: Chen, Kai-Shun
TITLE OF INVENTION: MAMMARY GLAND-SPECIFIC PROMOTERS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Charles & Brady
STREET: 411 East Wisconsin Avenue
CITY: Milwaukee
STATE: Wisconsin
COUNTRY: U.S.A.
ZIP: 53202-4497
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

```

[illegible]

OY 230 tccnagtacnaaggtcttcagac 253
Db 117 TCTCCATTACACGTGACCTAGGCC 94

RESULT 7
US-08-081-610-3
Sequence 3, Application US/08081610
Patent No. 5445941
GENERAL INFORMATION:
APPLICANT: Yang, Na N
TITLE OF INVENTION: Materials and Methods for Screening
- TITLE OF INVENTION: Anti-Osteoporosis or Serum Lipid Lowering Agents
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Allegretti and Witcoff, Ltd.
STREET: 10 S. Wacker Dr.
CITY: Chicago
STATE: IL
COUNTRY: U.S.A
ZIP: 60606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/081,610
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Heaphy, Barbara A
REGISTRATION NUMBER: 34,619
REFERENCE/DOCKET NUMBER: 93,402
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3303 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA

FEATURE:
NAME/KEY: mRNA
LOCATION: 2170..3303
FEATURE:
NAME/KEY: mRNA
LOCATION: 2214..3303
FEATURE:
NAME/KEY: mRNA
LOCATION: 2219..3303
FEATURE:
NAME/KEY: misc_RNA
LOCATION: 3301..3303
OTHER INFORMATION: /note= "CDS Start, codon start = 1,
OTHER INFORMATION: translation M"
FEATURE:
NAME/KEY: TATA_signal
LOCATION: 2170..2176
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1896..2306
OTHER INFORMATION: /note= "pb-301 -301 to +110"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1976..2306
OTHER INFORMATION: /note= "pb-221 -221 to +110"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2106..2306
OTHER INFORMATION: /note= "pb-91 -91 to +110"

FEATURE:
NAME/KEY: misc_feature
LOCATION: 2137..2306
OTHER INFORMATION: /note= "pb-60 -60 to +110"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2150..2306
OTHER INFORMATION: /note= "pb-47 -47 to +110"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2159..2306
OTHER INFORMATION: /note= "pb-38 -38 to +110"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2159..2271
OTHER INFORMATION: /note= "TGFB-3 position -38 to
OTHER INFORMATION: +75"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2159..2231
OTHER INFORMATION: /note= "TGFB-3 position -38 to
OTHER INFORMATION: +35"
US-08-081-610-3

Query Match 11.0%; Score 28; DB 1; Length 3303;
Best Local Similarity 59.0%; Pred. No. 2.2;
Matches 46; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

OY 170 gaaaggggtgaaggaaggggtaccttcagacacagaggaatgctgggga 229
Db 1229 GAAGAGGAGCGAAGACGACGACGCTTCCCTCACCATTCTGGGAATGCTCGGCC 1288

OY 230 tccnagtacnaaggtct 247
Db 1289 AACTGCTGACTTCGGCCT 1306

RESULT 8
US-08-687-559-2
Sequence 2, Application US/08687559
Patent No. 5955647
GENERAL INFORMATION:
APPLICANT: Fitcher, John H.
APPLICANT: Beachy, Roger N.
TITLE OF INVENTION: METHOD FOR USING TOBACCO MOSAIC VIRUS TO
- TITLE OF INVENTION: OVERPRODUCE PEPTIDES AND PROTEINS
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: US
ZIP: 92037

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/687,559
FILING DATE: No. 5955647ember 18, 1996
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/01467
FILING DATE: 03-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: learn, June M.
REGISTRATION NUMBER: 31,238
REFERENCE/DOCKET NUMBER: 07302/011001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-678-5070

MOLECULE TYPE: Genomic RNA

Oy 132 gaagatgtagtgaagaagccgnltggtgcTanccgatttgaaagaagtgaagaaaacgc 191
| ||||| | | ||||| | | | ||||| : | | ||||
Db 3059 GAGAGUGGUCGGCGGACGCCCGGUACAUCGANCUCACAAGCCCUUUGCAGNGGCACAGU 3118

OY 192 ggtaccttcagcacagcaggaatggtctg 225
Db 3119 CCUGACUUUUAACCAUCCGAAUAAAGAAGCUCUG 3152

RESULT 11

US-09-259-741-3
Sequence 3, Application US/09259741
Patent No. 6033895

GENERAL INFORMATION:

APPLICANT: GARGER, STEPHEN
APPLICANT: HOLTZ, R. BARRY
APPLICANT: MCCULLOCH, MICHAEL
APPLICANT: TURPEN, THOMAS

TITLE OF INVENTION: A PROCESS FOR ISOLATING AND
PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES FROM PLANT
SOURCES

CORRESPONDENCE ADDRESSES:

ADDRESS: Howrey & Simon
STREET: 1299 Pennsylvania Avenue N.W.
CITY: Washington
STATE: DC

COUNTRY: USA
ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/259,741
FILING DATE: February 25, 1999

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/037,751

FILING DATE: March 10, 1998
ATTORNEY/AGENT INFORMATION:

NAME: Halliuid, Albert P
REGISTRATION NUMBER: 25,277

REFERENCE/DOCKET NUMBER: 00801.0140.US01
TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-463-8100
TELEFAX: 650-463-8400

TELEX:

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:

LENGTH: 6425 base pairs
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: unknown

MOLECULE TYPE: Genomic RNA
US-09-259-741-3

Query Match 11.0%; Score 28; DB 3; Length 6425;
Best Local Similarity 46.8%; Pred. No. 3.1;
Matches 44; Conservative 8; Mismatches 42; Indels 0; Gaps 0;

OY 132 gaagatgagtggaagagccgntggctgctanccgattggaagagtggaagaaagcg 191
Db 3059 GGAAGAGUGUGCGGCGGAGCGCCGGAGUCCAGUCCGAAACCUCUUGCAUGGCAAGAU 3118

OY 192 ggtaccttcagcacagcaggaatggtctg 225
Db 3119 CCUGACUUUUAACCAUCCGAAUAAAGAAGCUCUG 3152

OY 192 ggtaccttcagcacagcaggaatggtctg 225
Db 3119 CCUGACUUUUAACCAUCCGAAUAAAGAAGCUCUG 3152

RESULT 12

US-09-037-751-3

Sequence 3, Application US/09037751
Patent No. 6037456

GENERAL INFORMATION:

APPLICANT: GARGER, STEPHEN
APPLICANT: HOLTZ, R. BARRY
APPLICANT: MCCULLOCH, MICHAEL
APPLICANT: TURPEN, THOMAS

TITLE OF INVENTION: A PROCESS FOR ISOLATING AND
PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES
FROM PLANT SOURCES

CORRESPONDENCE ADDRESSES:

ADDRESS: Howrey & Simon
STREET: 1299 Pennsylvania Avenue N.W.
CITY: Washington
STATE: DC

COUNTRY: USA
ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/037,751
FILING DATE: 10-MAR-1998

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER:

FILING DATE:
ATTORNEY/AGENT INFORMATION:

NAME: Halliuid, Albert P
REGISTRATION NUMBER: 25,277

REFERENCE/DOCKET NUMBER: 00801.0140.999
TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-463-8109
TELEFAX: 650-463-8400

TELEX:

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:

LENGTH: 6425 base pairs
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: unknown

MOLECULE TYPE: Genomic RNA
US-09-037-751-3

Query Match 11.0%; Score 28; DB 3; Length 6425;
Best Local Similarity 46.8%; Pred. No. 3.1;
Matches 44; Conservative 8; Mismatches 42; Indels 0; Gaps 0;

OY 132 gaagatgagtggaagagccgntggctgctanccgattggaagagtggaagaaagcg 191
Db 3059 GGAAGAGUGUGCGGCGGAGCGCCGGAGUCCAGUCCGAAACCUCUUGCAUGGCAAGAU 3118

OY 192 ggtaccttcagcacagcaggaatggtctg 225
Db 3119 CCUGACUUUUAACCAUCCGAAUAAAGAAGCUCUG 3152

OY 192 ggtaccttcagcacagcaggaatggtctg 225
Db 3119 CCUGACUUUUAACCAUCCGAAUAAAGAAGCUCUG 3152

RESULT 13

US-09-259-741-2

Sequence 2, Application US/09259741
Patent No. 6033895

GENERAL INFORMATION:

APPLICANT: GARGER, STEPHEN
APPLICANT: HOLTZ, R. BARRY
APPLICANT: MCCULLOCH, MICHAEL
APPLICANT: TURPEN, THOMAS

TITLE OF INVENTION: A PROCESS FOR ISOLATING AND
PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES FROM PLANT
SOURCES

CORRESPONDENCE ADDRESSES:

```

Query Match 11.0%; Score 28; DB 3; Length 6439;
Best Local Similarity 46.8%; Pred No. 3.1;
Matches 44; Conservative 8; Mismatches 42; Indels 0; Gaps 0;

OY 132 gaagatgagatgaagaagccgntgggtgctanccgattgaaagaagatgaagaagaagcg 191
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3059 GGAGAGUGCGCGCGGCGGCGCGCGGUGCAUACAUCGAUCUCAAACCCUUGCAUGCGAAGU 3118

OY 192 ggtacctcaagcacacagcgaggaatgctg 225
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3119 CCUGACUUUUVACCACAUCGGAUAGAAGAGCUCUG 3152

RESULT 15
US-09-259-741-5
: Sequence 5, Application US/09259741
: Patent No. 6038695

GENERAL INFORMATION:
APPLICANT: GARGER, STEPHEN
APPLICANT: HOLTZ, R. BARRY
APPLICANT: MCULLOCH, MICHAEL
APPLICANT: TURPEN, THOMAS
TITLE OF INVENTION: A PROCESS FOR ISOLATING AND
TITLE OF INVENTION: PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES FROM PLANT
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESS: Howrey & Simon
STREET: 1299 Pennsylvania Avenue N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20004

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/259, 741
FILING DATE: February 25, 1999
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/037, 751
FILING DATE: March 10, 1998
ATTORNEY/AGENT INFORMATION:
NAME: Halluin, Albert P

```

```

: REGISTRATION NUMBER: 25,277
: REFERENCE/DOCKET NUMBER: 00801.0140.US01
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 650-463-8100
: TELEFAX: 650-463-8400
: TELEX:
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 6446 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: unknown
: MOLECULE TYPE: Genomic RNA
: US-09-259-741-5

```

```

Query Match      11.0%; Score 28; DB 3; Length 6446;
Best Local Similarity 46.8%; Pred. No. 3.1;
Matches 44; Conservative 8; Mismatches 42; Indels 0; Gaps 0;

132 gaagatgagatgaaagagcagntgggtgctanccgattgaaagagatgaaagaaagcg 191
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3059 GGAGAUUGGUGCGGCGAGCGCGCGGUGAUCUAUCCGAUCUAACCAACCCUUGCAUGGCAAGAU 3118
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 192 ggtatctcaagcacacagcaggaatggtctg 225
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 3119 CCUGACUUUUUACCAUCCGAUAAAGAACUCUCG 3152
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: March 18, 2001, 08:26:06
 Job time: 3808 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 18, 2001, 07:17:18 ; Search time 954.67 Seconds
(without alignments)
1361.631 Million cell updates/sec

Title: US-09-199-129A-1
Perfect score: 254
Sequence: 1 ggaatctgaagctaccta.....atgacnaagctctcagaca 254

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 118133 seqs, 2558875100 residues

number of hits satisfying chosen parameters: 2236266

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba1: *
2: gb_ba2: *
3: gb_om: *
4: gb_om: *
5: gb_ph: *
6: gb_p11: *
7: gb_p12: *
8: gb_p13: *
9: gb_p14: *
10: gb_p15: *
11: gb_p16: *
12: gb_p17: *
13: gb_p18: *
14: gb_p19: *
15: em_fun: *
16: em_hum1: *
17: em_hum2: *
18: em_hum3: *
19: em_hum4: *
20: em_hum5: *
21: em_hum6: *
22: em_hum7: *
23: em_hum8: *
24: em_hum9: *
25: em_hum10: *
26: em_hum11: *
27: em_hum12: *
28: em_hum13: *
29: em_hum14: *
30: gb_in1: *
31: gb_in2: *
32: gb_in3: *
33: gb_in4: *
34: gb_in5: *
35: gb_in6: *
36: gb_in7: *
37: gb_in8: *
38: gb_in9: *
39: gb_in10: *
40: gb_in11: *
41: gb_in12: *
42: gb_in13: *
43: gb_in14: *

44: em_hcg8: *
45: em_hcg9: *
46: em_hcg10: *
47: em_hcg11: *
48: em_hcg12: *
49: em_hcg13: *
50: em_hcg14: *
51: em_hcg15: *
52: em_hcg16: *
53: em_hcg17: *
54: em_hcg18: *
55: em_hcg19: *
56: em_hcg20: *
57: em_hcg21: *
58: em_hcg22: *
59: em_hcg23: *
60: em_hcg24: *
61: em_hcg25: *
62: em_hcg26: *
63: em_hcg27: *
64: em_hcg28: *
65: em_hcg29: *
66: em_hcg30: *
67: em_hcg31: *
68: em_hcg32: *
69: em_hcg33: *
70: em_hcg34: *
71: em_hcg35: *
72: em_hcg36: *
73: em_hcg37: *
74: em_hcg38: *
75: em_hcg39: *
76: em_hcg40: *
77: em_hcg41: *
78: em_hcg42: *
79: em_hcg43: *
80: em_hcg44: *
81: em_hcg45: *
82: em_hcg46: *
83: em_hcg47: *
84: em_hcg48: *
85: em_hcg49: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	95.6	37.6	1424	7	ATU66345
2	77	30.3	1564	7	BVNRNAC
3	76.4	30.1	1545	6	AF190454
4	76.2	30.0	1463	7	ATU66343
5	75.6	29.8	1547	6	AF052040
6	75.6	29.8	1593	7	AF283816
7	74.4	29.3	1558	33	NPCALIMNR
8	74	29.1	1450	7	BLYCRH2A
9	74	29.1	1472	7	BLYCRH1A
10	74	29.1	1558	6	AB021259
11	72.6	28.6	1514	33	RCU74630
12	72.6	28.6	1543	7	CAPCRIC
13	72	28.3	89479	6	AC006932
14	71.6	28.2	1614	33	ZMCALPREC
15	70	27.6	1342	33	ZMCRTIGEN
16	70	27.6	1413	7	ATU27698
17	69.4	27.3	1524	6	AF134733
18	66.8	26.3	1170	33	NTRNATCAL
19	65.6	25.8	262	78	RICL819A
20	65.2	25.7	1600	6	AF019376
21	56.2	22.1	1363	32	DDU36937

22	54.2	21.3	98616	67	AC058785	Arabidops
23	54.2	21.3	112448	69	AC069159	Arabidops
24	53.4	21.0	2038	7	AT066344	Arabidops
25	53.4	21.0	59261	33	TI1244	Arabidops
26	51.4	20.2	49275	33	RCU74631	Arabidops
27	43.6	17.2	732	6	AB018243	Arabidops
28	43.6	17.2	2336	7	CRAJ765	Arabidops
29	37	14.6	98616	67	AC058785	Arabidops
30	35	13.8	195154	62	AC022246	Arabidops
31	35	13.8	276042	57	AC012183	Arabidops
32	34.6	13.6	25842	31	CEP57F5	Homo sapi
33	34.6	13.6	159639	55	AC008747	Caenorhabd
34	34.6	13.6	160026	61	AC021437	Homo sapi
35	34.4	13.5	136519	66	AC0344229	Homo sapi
36	34.2	13.5	95683	57	AC011796	Homo sapi
37	34.2	13.5	205839	57	AC013350	Homo sapi
38	34	13.4	109614	68	AC064850	Mus musc
39	34	13.4	183628	57	AC012450	Homo sapi
40	33.8	13.3	43345	58	AC014886	Homo sapi
41	33.8	13.3	80074	9	AC008389	Drosophila
42	33.8	13.3	84970	54	AC008389	Homo sapi
43	33.8	13.3	86581	59	AC016426	Homo sapi
44	33.8	13.3	101076	52	HS37M17	Homo sapi
45	33.8	13.3	106926	54	AC007691	Human DNA s
						AC007691 Drosophila

ALIGNMENTS

[illegible]

RESULT	1
LOCUS	ATU66345
DEFINITION	ATU66345
ACCESSION	U66345
VERSION	U66345.1
KEYWORDS	GI:2052382
SOURCE	thale cress.
ORGANISM	Arabidopsis thaliana
REFERENCE	Eukaryote: Viridiplantae: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots; Rosidae; eurosids II: Brassicales; Brassicaceae; Arabidopsis.
AUTHORS	1 (bases 1 to 1424)
TITLE	Nelson,D.E., Glaunsinger,B. and Bohnerl,H.J.
JOURNAL	abundant accumulation of the calcium-binding molecular chaperone calreticulin in specific floral tissues of Arabidopsis thaliana
MEDLINE	Plant Physiol. 114 (1), 29-37 (1997)
WORDS	97303616
REFERENCE	2 (bases 1 to 1424)
WORDS	Nelson,D.E., Glaunsinger,B. and Bohnerl,H.J.
JOURNAL	Direct Submission
WORDS	Submitted (07-AUG-1996) Biochemistry, University of Arizona,
REFERENCE	Bioscience West 513, Tucson, AZ 85721, USA
WORDS	Location/Qualifiers
FEATURES	1..1424
SOURCE	/organism="Arabidopsis thaliana"
gene	/db_xref="taxon:3702"
CDS	23..1297
	/gene="Crt3"
	23..1297
	/gene="Crt3"
	/codon_start=1
	/product="calreticulin"
	/protein_id="AAC49697.1"
	/db_xref="GI:2052383"
	/translation="MGLPQNKLSFECFFLVLTAPLAFSEITLEHREGGKSRM
	VLSDDKRNKERAGTEFKHTPAKWPDPDNKIQYNDKAKVAISAKIPEFSKNKNTLVV
	QYSAKLEODIECGGAYIKLISGYVNGQFGSDPYSLMFGDICTQTKLHVIVSYQ
	GQNPVPIKDDCETCDLNHYETFLRPDASVYLNDKREFGSGMTDMDLPPRKIK
	VKNAKRGEKMDREYIDDPNDVMEGDSIPRELPDKAKPEMDDEENGILAMPKIK
	PNVAYGPKAKRKIKNNPNYGMKNKMPIDNPEPDDPDLYIKSITKAGIVNWKAG
	SIDNLTICDDPAAVARSIVDDYPAQHESKEKFLPAEKERKAKNEDEARAKREEGE
	RRKEDRIKDRKRIKRPMDIMDYHDEL"
BASE COUNT	449 a 250 c 350 g 360 t

[illegible]

Oy	20	ATGGAGAAATCGGCGGCGGAACCCCTAGCTTTTCTCTCTCCTTCCTCTTCTTTT	79
Oy	93	cactcttcattcgtcgagatcttcttcgaagaagaaatccgaagtgaaggacgcgn	152
Db	80	GCCATCATGTCTGTAAGTGTTCTTCAGAGAAGCTTTTCGAAGATGATGGGAAGCGT	139
Oy	153	tgggtgctancncgatctggaagaagagtgaaagaaaagcggtacctccaagcacagca	212
Db	140	TGGGTCAATCTGATGTGAGAGAAGATGAAATATGCGTGGAGAGATTGAAATTACATCC	199
Oy	213	ggaanaatggctcgggataccngaatgacnaagttccagac	253
Db	200	GCCTAATGGAACGAGATGCCAATGACAACAGTATCCAGAC	240
RESULT	3		
AFI90454			
LOCUS	AFI90454	1545 bp	mRNA
DEFINITION	Zea mays calreticulin (CRT) mRNA, complete cds.		PLN
ACCESSION	AFI90454		
VERSION	AFI90454.1	GI:6014706	
SOURCE			
ORGANISM	Zea mays.		
REFERENCE	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.		
AUTHORS	Wyllie, W.E., Tsou, P.-L. and Robertson, D.		
TITLE	Effects of altered expression of the calcium-binding protein calreticulin in Arabidopsis thaliana		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1545)		
AUTHORS	Wyllie, S.E., Mrobel, R., Boston, R. and Robertson, D.		
TITLE	Direct Submission		
JOURNL	Submitted (27-SEP-1999) Botany, North Carolina State University,		
FEATURES	Box 7612, Raleigh, NC 27695, USA.		
source			
gene			
CDS			
ORIGIN			
BASE COUNT	454 A 339 C 423 G 329 T		
Query Match	30.1%; Score 76.4; DB 6; Length 1545;		
Best Local Similarity	69.2%; Pred. No. 1;le-13;		
Matches 101: Conservative	2% Mismatches 45; Indels 0; Gaps 0;		
Oy	108	gagattctcttcgaagaagatctgaagaatgtagaagagcggttggtgtctanccgat	167
Db	153	GAGGCTCTTCTTCACGACAGAACATGCGTGGGAAAGTCGCTGCAAGTCCAG	212
Oy	168	tggaaaaagagtgaaagaaacggttaacctcaagcacacagcaagaaatlggtctgg	227
Db	213	TGCAAGAACGATGAGAACATGCTGCTGAATGGAACCACACATCTGGAATAATGGAATGCA	272

Oy	228		gatccnagtcgaaggctcaccgcacg	253
Dd	273	GATCGCAGGACAAAGATTCAAAC	298	

RESULT 4
 ATU66343
 LOCUS Arabidopsis thaliana calreticulin (Crcl) mRNA, complete cds.
 DEFINITION Accession U66343
 VERSION U66343.1 GI:2052378
 KEYWORDS
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 Magnoliophyta; eudicotyledons; core eudicrots; Rosidae; eurosids II;
 Brassicales; Brassicaceae; Arabidops.

REFERENCE
 AUTHORS Nelson,D.E., Glanunsinger,B. and Bohnert,H.J.
 TITLE Abundant accumulation of the calcium-binding molecular chaperone calreticulin in specific floral tissues of Arabidopsis thallana Plant Physiol. 114 (1), 29-37 (1997)
 JOURNAL MEDLINE 97303616
 REFERENCE 2 (bases 1 to 1463)

AUTHORS Nelson,D.E., Glannusinger,B. and Bohnert,H.J.
 DIRECT SUBMISSION Submitted (07-AUG-1996) Biochemistry, University of Arizona,
 Biosciences West 513, Tucson, AZ 85721, USA

JOURNAL LOCATION/Qualifiers
 TITLE Location/Qualifiers
 FEATURES
 source 1..1463 /organism="Arabidopsis thaliana"
 /gene "crcl"/db_xref="taxon:3702"
 4..1281 /gene="Crcl"
 4..1281 /codon_start=1
 /codon_start=-1
 /product="calreticulin"
 /protein_id="AAC49695.1"
 /db_xref="GI:2052379"

CDS
 1 translation="MAKLNPETSLIFALVIVISAEVIPEEKPEDMGEMKRWKSDMK
 KDNTTAGEKKHTTAGNMSGDANDKIQTISEDRFAISAEPFESNKKTILVFQSYVKR
 EOKLDCGGGYMGLSDVDVGRTPEGDPYSTIMEGDYSTKVVAHLYTGNNHLI
 KEPCEFTDLTHVTFLVRPDADAYSLIDNEVOTSGLSDMWLPAKITPSARAK
 PEMDDEKEYIPDEDTKTPAGYCDDLPKIKPTDAKPEDWDDEEGEWTAPIPMPEYN
 GEMLPKINPVAVYGKMKAVIDNPERKDDPELVPRKLKYGVELMOVGSGLFDNV
 LVSDPEYAKIAEETKGKHDKAENAFDEAKEERRESEESDNAESADEEAADDONN
 EGGSDNESSESSETRKEARETEFKAEETAHADEL".
 BASE COUNT 456 a 279 c 368 g 360 t
 ORIGIN

```

Query Match          30.0% Score 76.2 ; DB 7 ; Length 1463 :
Best Local Similarity 68.5% Pred. NO. 1.3e+15 ;
Matches 102; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
```

Oy	105	gccagaatttcttcgaagaagatcgcaagaagtgaagaagccgnltgggtcanc	164
Dd	67	GCTGAAGTTACTTCGCAGAGAANAATCCAAAGATGTGGAGAAGAGATGGATAAAATCT	126
Oy	165	gatcgaaaaggagtgaaagaaaagcgggtactccctcaaagcacacagacagaaataatgctct	224
Dd	127	GACTGCCAAGAAAAGATTAATACTGCTGGGGAATGGAACAACACACGCTGAATAATTGGTCT	186
Oy	225	ggggatccngatgcnaaatgctccaagc	253
Dd	187	GGTGATGCTAACGANTAAAGGTATCCAAC	215

RESULT 5	AF052040	1547 bp	mRNA	PLN	04-MAR-1999
LOCUS	DEFINITION	Berberis stolonifera calreticulin mRNA, complete cds.			

ACCESSION AF052040 GI:4335863
VERSION
KEYWORDS Berberis stolonifera.
SOURCE Berberis stolonifera
ORGANISM Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; Ranunculales; Berberidaceae;
Berberis.

REFERENCE 1 (bases 1 to 1547)
AUTHORS Chou,W.-M. and Kutchan,T.M.
TITLE Calreticulin from Berberis stolonifera
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1547)
AUTHORS Chou,W.-M. and Kutchan,T.M.
TITLE Direct Submission
JOURNAL Submitted (03-MAR-1998) Laboratorium fuer Molekulare Biologie,
Universitaet Muenchen, Karlstrasse 29, Muenchen 80333, Germany

FEATURES
source location/Qualifiers
1..1547
/organism="Berberis stolonifera"
/db_xref="taxon:33814"
/note="Cell suspension culture"
21..1271
/function="possible molecular chaperon"
/note="calcium binding protein of the endoplasmic reticulum"
/codon_start=1
/product="calreticulin"
/protein_id="AAI17490.1"
/db_xref="GI:4335864"
/translation="MAIAERSRSHLARVRDRVSAEYFEFEREDGESKVKSDMKR
RDENWAGEMNFTSGKMNGDANDKCIQTSEDYRFYAISAPEFSNKETLVEFQSVKK
EQKLDGCGGYMKLLSGVDQKRFSGDPYSIMFPDLCGVSTKRVAHLITGEINHLIL
KQIDPCEDTDLTHYYTFELRPDAVSILLIDVESGSYGVDMDILPPKOIKDEPAEK
PEDMEDKEVIPDEDKPEGDDIPKETIDEAKPEDEMDEEGEMVTAPTPNDYVK
GEWKPKIKNPNEFDKGKKAPMIQNDPKDDPIYVPKIKYVGTEIQVKSQSYTFDNV
LICDDPYRAKLAEETGWCKNDADKAFADEAEKKKEEBEAKDPTESDDEKPDDEGESE
DGECDDSKDINDNEDEDVHDEL"

BASE COUNT 487 A 277 C 391 G 392 T
ORIGIN

Query Match 29.8%; Score 75.6; DB 6; Length 1547;
Best Local Similarity 62.0%; Pred. No. 2e-13;
Matches 114; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

OY 70 ctgagcttcttcnaactcgcacctcgaacctcaatcgccggagagtcttcgaagaagagat 129
|| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
F 49 CTCATCTTGCMCTCCGCTTTGGCAGTGCCTGTTCCTTCGCGAAGCTCTTTCGAAGAAGCAT 108
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 130 tcgaagatgata tggaaagaccgnltbgtygtctancaccgatgtgaaaaggatgaagaagaaag 189
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 109 TTGAAGATGATGGAGAGCAATAATGGTTAATCATCGATTGGAAGAGGATGTGAACAATCG 168
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 190 cgggtacctaagcacacagcagcaaaaaatgtycttggaattccngatgacmaaggtcttc 249
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 169 CCGGAGACTGGAACCTTCACTCCGGAATAATGGAATGAGATGATGATAAAGGTATCC 228
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 250 agac 253
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 229 AAAC 232

RESULT 6
AF283816 1593 bp mRNA PLN 21-AUG-2000
LOCUS
DEFINITION Pinus taeda calreticulin mRNA, complete cds.
ACCESSION AF283816
VERSION
KEYWORDS
SOURCE loblolly pine.
ORGANISM Pinus taeda
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	FEATURES
1	Coniferopsida: Coniferales; Pinaceae; Pinus.	1 (bases 1 to 1593)					
2	Shealy, M.J. and Frankis, R.C. Jr.	Complete nucleotide sequence of a cDNA encoding calreticulin from Pinus taeda	Unpublished				
3	2 (bases 1 to 1593)						
4	Shealy, M.J. and Frankis, R.C. Jr.	Direct Submission					
5	Submitted (30-JUN-2000) Department of Biology, College of Charleston, 66 George St., Charleston, SC 29424, USA						
6	Location/Qualifiers						
7	1..1593						
8	/organism="Pinus taeda"						
9	/db_xref="taxon:3552"						
10	/tissue="type="pollen"						
11	24..1307						
12	/note="calcium-binding protein"						
13	/codon_start=1						
14	/product="calreticulin"						
15	/protein_id="PAG01147.1"						
16	/db_xref="GI:9858814"						
17	/translation="MAGRRSLVAVFLLEPTLVSAEYFEFEREDDSMESRWOSDMKDESLAGPMVHTSGKMGNDPMDKGIOTHTDVRFAISAAYEPFNKKTIVLPLOSVRKESQKIDCGGYYKILSGEIDKQNFSGEPYSTMFEPRDTCSTKKYHTILSKGNPIKDYCEEDQLTHVTFILRPDATTSLINDNDKOSGLIKWDLLPRTIKDPNPKKEREDWDRKEYTDPEDEKPREGYDDIPEKIDPDATPEWMDNDEEWAPITANPEYKGMKWRKIKINPEYKRWAPMIDNDFKQDPELVYFNLKYGILMAYKSGTLFDNIISDPEYAKKILAEYTMARHNDKAEKAEFADEAKKKKEEKESDDEEDTDEKESDDEADDELDEHKEADKKHEHLEDEHKKHDEHDEL"						
18	BASE COUNT	532 a	267 c	382 g	412 t		
19	ORIGIN						
20	Query Match	29.8%; Score 75.6; DB 7; Length 1593;					
21	Best Local Similarity	68.0%; Pred. No. 2.1e-13;					
22	Matches 102; Conservative	0; Mismatches 48; Indels 0; Gaps 0;					
23	OY	105 gctgagattcttcctgcgaagaagatcgcgaatggaatggaagcgcgntggtctancc 164					
24	DB	87 GCCGAAGTCTTTTTCGAGAGAGCGTTTCGATTCAGTCTGGGAAAGTCGTTGGTGCATCT 146					
25	OY	165 gattggaagaagagatggaagaagcgggttaccttcacgacacacgcaggaataatgctc 224					
26	DB	147 GATTGGAAAAGAGATGAAGACTTGCTGAGATTGGGTGCACACATCAGGGAAATGGAAC 206					
27	OY	225 ggggagtcnagatcnaagatcttcacaca 254					
28	DB	207 GGGGATCCCAATGACACAGGCGCATCCACAACA 236					
29	RESULT	7					
30	NCPCALIMNR	1558 bp	mRNA	PLN	27-JUN-1996		
31	LOCUS	N.Piunbagnifolia mRNA for calreticulin.					
32	DEFINITION	271395					
33	VERSION	271395.1	GI:1419087				
34	KEYWORDS	CAL1 gene; calreticulin.					
35	SOURCE	curled-leaved tobacco.					
36	ORGANISM	Nicotiana glauca					
37	REFERENCE	1					
38	AUTHORS	Borisjuk, N., Borisjuk, L., Adler, K., Stailo, L., Tewes, A. and Martenfele, R.					
39	TITLE	Differential expression of calreticulin during somatic and zygotic embryogenesis of Nicotiana glauca					
40	JOURNAL	embryogenesis of Nicotiana glauca					
41	REMARK	(sites)					
42	REFERENCE	2 (bases 1 to 1558)					
43	AUTHORS	Borisjuk, N.					
44	TITLE	Direct Submission					


```

/gene="CRH1"
/codon_start=1
/function="calcium binding protein"
/product="calreticulin"
/protein_id="AA32948.1"
/db_xref="GI:439586"
/translation="LLRRLLALLASVAANAADVFQDEKPEDGSESNVWVSEMKKDEEN
MAGWNHTSGKRWGHDAEDKGIQTSSEDRFAISAEVPEFSKKDITLAIQFVKEOKL
DCGGGYKLLGLGDVDOKKFGSDPRPYGIMFSDICGYSIKVHTLITLNGKNHLLIKK
PCEDDLSHYTLLIRPDATYSLIDNEKOTGSIYHMWLLPKRLIKDPAKKPEDM
DKKEYIPDPEDVKEGEGYDDIPKEVTPDDAKKPEDMDDEDEGEMTAPITIPNEYKPMKK
OKRTKKNPYGKKWAPMIANDPFQDDYIYAFDSLKTYGLELWQVKSGLTFDNLITVD
DAALIAFAETWAKHKDAEKAADFDEAEKKKEEDDASKAGEDDDDLDEADDEKDDK
KAGSDAEDDKDSDDEKDEEL"

```

	Query Match	29.1%	Score 74	DB: 74	Length 1472
	1st Local Similarity	67.3%	Pred. No. 6.6e-13		
	Matches 101	Conservative 0	Mismatches 49	Indels 0	Gaps 0
QY	104	cgcttgaagattctcttcgaaagagagattcgaagatgtagaagacgcagtgglyctanc	163		
Db	53	CGCGACGCTCTTCCTCCAGGAGAGATTGCAAGATGCGCTGGGAACCGGTCGTCAAATC	112		
QY	164	cgattggaagaagagtggaagaaacggggtacccttcaagcacacagcaggaanaatgctc	223		
Db	113	TGAGTGGAGGAAGGACGCGAACAATGGCGTGTGAATGAGAACACACATCTGGAAATATGCCA	172		
QY	224	tggggattcgnatgacnaaagattcttcagac	253		
Db	173	TGGAGATCTCTGAGGACANAGGATATCCAAAC	202		

```

RESULT 10
AB021259
LOCUS      AB021259      1558 bp      mRNA      PLN      28-MAR-2000
DEFINITION Oryza sativa mRNA for calcium-binding protein, complete cds.
ACCESSION  AB021259
VERSION    AB021259.1 GI:6682832
KEYWORDS   calcium-binding protein.
SOURCE     Oryza sativa cDNA to mRNA.
ORGANISM   Oryza sativa
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.
1 (sites)
Li, Z. and Komatsu, S.
Molecular cloning and characterization of calreticulin, a
calcium-binding protein involved in the regeneration of rice
cultured suspension cells
Eur. J. Biochem. 267 (3), 737-745 (2000)
2011/682
2 (bases 1 to 1558)
Komatsu, S. and Li, Z.
Direct Submission
Submitted (15-DEC-1998) to the DDBJ/EMBL/genbank databases.
Setsuko Komatsu, National Institute Agrobiological Resources, Department of
Molecular Biology, Kannondai 2-1-2, Tsukuba 305-8602, Japan
(E-mail:skomatsu@agr.affrc.go.jp, Tel:81-298-38-7490,
Fax:81-298-38-7032)
1558
Location/Qualifiers
1..1558
/organism="Oryza sativa"
/db_xref="taxon:4530"
31..1305

```

```

/codon_start=1
/product="calcium-binding protein"
/protein_id="BA88900.1"
/db_xref="GI:6682833"
/translation="MAIRARSSSSAAAAVALALALAAVAAVAGVEVFQOEKFEDEMER
WYSEKKEDENMACEMNHTSGKMGDEDEDGIQISEYRFAISAETPERSNKITVL
LQSVKREKQLDCGGGYVLLGGGVDDKRGCGDIPYAGCPDLCGSTRKVVHTFPR

```

BASE COUNT 453 a 330 c 437 g 337 t 1 others
ORIGIN

Query Match	Score	DB	Length
Best Local Similarity	67.3%		
Matches	101	Conservative	0
		Mismatches	49
		Indels	0
		Gaps	0

RESULT	11
LOCUS	RCU74630
DEFINITION	RcU74630 1514 bp mRNA PLN 29-SEP-1997
ACCESSION	U74630
VERSION	U74630.1 GI:1658196
KEYWORDS	.
SOURCE	castor bean.
ORGANISM	Ricinus communis
REFERENCE	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; Rosidae; eurosids I; Malpighiales; Euphorbiaceae; Ricinus.
REFERENCE TITLE	1 (bases 1 to 1514) Coughlan,S.J., Hastings,C. and Winfrey,R. Jr. Cloning and characterization of the calreticulin gene from Ricinus communis L Plant Mol. Biol. 34 (6), 897-911 (1997)
JOURNAL MEDLINE	97435975
REFERENCE	2 (bases 1 to 1514)
AUTHORS	Coughlan,S.J., Hastings,C. and Winfrey,R.J.
TITLE	Direct Submission
JOURNAL	Submitted (15-OCT-1996) TetD, Pioneer Hi-Bred, 7300 NW 62nd Avenue, Johnston, IA 50131, USA
FEATURES	Location/Qualifiers
source	I..1514

```

CD5

/organism="Ricinus communis"
/db_xref="taxon:3988"
7. .1254
/codon_start=1
/product="calreticulin"
/protein_id="AA071419.1"
/db_xref="GI:1658197"
/translation="MANPKSLFLSLTSLAASAEVFEERFEDGEMNRYKSDMKDD
ENRAGEWNTYSGKMGNDPNKGTQTSIEDRYEVAISAEFFESKDKTIVYQPSVKHKO
KLDCGGCYMKLLSSSTDRKKFEGDTPISMGFDICGSTRKHYHALINYNDTNHLIK
EVEPEYDQLTHVTVLTVPATYSILDNVEKOTGSLYTDWDLPPKIKDPEAKPE
DMDEKEYIDPEDRKKEGVDIDPRELEIDPPAKRPEWDDEDEDEWPTAIPEYKGP
WKPKIINPNYKGMKAPMIDNDFKDDPEIYYPNLKTYGIELMGVKSGLPDNYLI
CNDEYAKOLAEETWCKMGKNDPAEKAAAEPAKKKEEESKDDPADSDAEDDDADDE
GDDGEEKSNAADSAEDVDHEL"

BASE COUNT      505 a      268 c      363 g      378 t
ORIGIN

```

Query Match	28.6%	Score 72.6;	DB 33;	Length 1514;
Best Local Similarity	65.88;	Pred. NO. 1.9e-12;		
Matches 102; Conservative	0;	Mismatches 53;	Indels 0;	Gaps 0;

QY	100	catctgcgcgaagatctctctctcgaagaagatctcgaagaatgaggaagccgntggctgc	159
Db	59	CTTCTGCGAAGACTCTTCTTCCGAGAGCGCTTTGCAAGATGATGCGAAATAGACTGGCTTA	118
QY	160	tanccagatctgaaagaagatgaaagaaacgagtaaccttcagacacacagcaggaanaat	219
Db	119	AATCTGATGGAAGAAAGATGAGATACAGCTGGTGAATGAAATATATACCTCTGGAAAGT	178
QY	220	ggctcgggagatcngatgacmaaggctctcgaaca	254
Db	179	GGAATGGAGACCCCTAATGACAAAGGATATTCAACA	213

RESULT	12			
CAPCRIC	1543 bp	mRNA	PLN	08-SEP-1995
DEFINITION	C. annum PCRIC mRNA.			
ACCESSION	X80756			
VERSION	X80756.1 GI:984112			
WORDS	calreticulin.			
ORIGIN	Capsicum annum.			
ORGANISM	Capsicum annum.			
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; eudicotyledons; Spermatophyta; Magnoliophyta; eudicotyledons; Asceridae; Solanaceae; Solanales; Solanaceae; Capsicum.			
AUTHORS	Hugueney,P., Bouvier,F., Badillo,A., d'Harlingue,A., Kuntz,M. and Camara,B.			
TITLE	Identification of a plastid protein involved in vesicle fusion and/or membrane protein translocation			
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 92 (12), 5630-5634 (1995)			
MEDLINE	95296365			
REFERENCE	2 (bases 1 to 1543)			
AUTHORS	Camara,B.			
TITLE	Direct Submission			
JOURNAL	Submitted (29-JUL-1994) B. Camara, IBMP CNRS, 12 rue du General Zimmer, 67084 Strasbourg Cedex, FRANCE			
FEATURES	location/Qualifiers			
source	1..1543			
BASE COUNT	452 a 312 c 376 g 403 t			
ORIGIN	/organism="Capsicum annum" /db_xref="taxon:4072"			

Query Match	28.6%	Score 72.6	DB 7	Length 1543
Best Local Similarity	59.4%	Pred. No. 1.9e-12		
Matches 117	Conservative 0	Mismatches 80	Indels	Gaps 0
QY	57	tcgctcaagcactctgcttctctctcnaactcgcgaactcttcagctcagatcttc	116	
Db	1	TCTCTCTATCTCATTCGCTGCTTTCTCTCTACTTGTAGCTTCCGCCGCCCAAGATTTC	60	
QY	117	ttcgaagaagagatctcgaagaatgaggaagccgntggctcancagatctgaaag	176	
Db	61	TTTCGAGGAAATTTTCAATGATGCTGTGGAGACACAGATGGGTGAATCTCATTTGGAAGAA	120	
QY	177	agtgagaagaagacgggtctaccttcacgaacacagcaggaanaatgctctgggagatcngat	236	
Db	121	GACGAGAACATGTGCGCGAGATGTGAATCACACCTCTGGCMAAGTGGAGTGAAGACGCTAAT	180	
QY	237	gacnaaggtctctcagac	253	
Db	181	GACAAAGCTATCCACAC	197	

RESULT	13			
AC006932	89479 bp	DNA	PLN	11-OCT-2000
LOCUS	AC006932/c			
DEFINITION	Genomic sequence for Arabidopsis thaliana BAC T2767 from chromosome I, complete sequence.			
ACCESSION	AC006932			

VERSION	AC006932.8	GI:6468048
KEYWORDS	HTG.	
SOURCE	thale cress.	
ORGANISM	Arabidopsis thaliana	
REFERENCE	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.	
AUTHORS	1 (bases 1 to 89479) Shinn,P., Brooks,S., Buehler,E., Chao,Q., Johnson-Hopson,C., Khan,S., Kim,C., Altafi,H., Bel,O., Chin,C., Chlou,J., Choi,E., Conn,L., Conway,A., Gonzales,A., Hansen,N., Howing,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,J., Liu,K., Liu,S., Mukharsky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A., Thaveli,A., Toriumi,M., Vaysberg,M., Yu,G., Federspiel,N.A., Theologis,A. and Ecker,J.R.	
TITLE	Genomic sequence for Arabidopsis thaliana BAC T2767 from chromosome I	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 89479) Ecker,J.R.	
AUTHORS	Direct Submission	
TITLE	Submitted (02-MAR-1999) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA	
JOURNAL	3 (bases 1 to 89479) Ecker,J.R.	
REFERENCE	Direct Submission	
AUTHORS	Submitted (20-MAY-1999) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA	
TITLE	4 (bases 1 to 89479) Ecker,J.R.	
REFERENCE	Direct Submission	
AUTHORS	Submitted (30-OCT-1999) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA	
JOURNAL	5 (bases 1 to 89479) Ecker,J.R.	
REFERENCE	Direct Submission	
AUTHORS	Submitted (25-NOV-1999) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA	
TITLE	6 (bases 1 to 89479) Chao,Q., Brooks,S., Buehler,E., Johnson-Hopson,C., Khan,S., Kim,C., Shinn,P., Altafi,H., Bel,B., Chin,C., Chlou,J., Choi,E., Conn,L., Conway,A., Gonzales,A., Hansen,N., Howing,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,J., Liu,S., Mukharsky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A., Thaveli,A., Toriumi,M., Vaysberg,M., Yu,G., Davis,R., Federspiel,N., Theologis,A. and Ecker,J.	
JOURNAL	Direct Submission	
REFERENCE	Submitted (03-JAN-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA	
AUTHORS	7 (bases 1 to 89479) Chao,Q., Brooks,S., Buehler,E., Johnson-Hopson,C., Khan,S., Kim,C., Shinn,P., Altafi,H., Bel,B., Chin,C., Chlou,J., Choi,E., Conn,L., Conway,A., Gonzales,A., Hansen,N., Howing,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,J., Liu,S., Mukharsky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A., Thaveli,A., Toriumi,M., Vaysberg,M., Yu,G., Davis,R., Federspiel,N., Theologis,A. and Ecker,J.	
TITLE	Direct Submission	
JOURNAL	Submitted (11-OCT-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA	
COMMENT	On Nov 25, 1999 this sequence version replaced gi:6143825.	
FEATURES	Location/Qualifiers	
SOURCE	1..89479 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /chromosome="1" /clone="T2767"	

FEATURES	source
REFERENCE	1 (bases 1 to 1614)
AUTHORS	Dresselhaus,T., Hagel,C., Lott,H. and Kranz,E.
TITLE	Isolation of a full-length cDNA encoding calreticulin from a PCR library of in vitro zygotes of maize
JOURNAL	Plant Mol. Biol. 31 (1), 23-34 (1996)
MEDLINE	96309381
REFERENCE	2 (bases 1 to 1614)
AUTHORS	Dresselhaus,T.
TITLE	Direct Submission
JOURNAL	Submitted (19-JUL-1995) T. Dresselhaus, Applied Plant Molecular Biology II, University of Hamburg, Ohnhorststr. 18, D-22609 Hamburg, FRG
FEATURES	Location/Qualifiers
source	1. 1614
	/organism="Zea mays"
	/cultivar="inbred line A188"
	/db_xref="taxon:4577"
	/dev_stage="in vitro zygotes, 18h after in vitro fertilisation"
	/clone_lib="cDNA library of zygotes"
	/clone="ZM2 233"
	76. 1341
	/codon_start=1
	/product="calreticulin precursor"
	/protein_id="CA61939.1"
	/db_xref="GI:927572"
	/db_xref="SPRMBL:O43712"
	/translation="MAIRKSSVAVALALALASVAAGVEYFOEKEDGSESRVSKS EMKEDNAGCEMNTSGKMNDAEDKQGTSEDRFYIASNEYEFSSKQTVLYQTS VKHEOKLDGGGYVKLLAGDVQDKFGSDISYSAFPGDTCYSTKRVHTLLYDGRKK HLIKRDVPCETDQLTHTVYTLIRPDATYSILIDNEKQTSIYEHMDLVPKRTKDEE AKKEDMDDEKYEIPDEDKRPEGVDLPKEIIPDDAKPEDMDDEDEDEWTAPIPNNE EYKPMKOKRIKKNFYOCWKAKAPMIDNDDPEDDPYIYAFSLKTYIGTELMOVKSGLLE DNITITDPAIAKTPAEFTWCKHKKEAEKAAPDEAKKKEEDDAKKGDEDDDLDEDE DDEKADKADSDAEDSKSDDEKQHDEL"
polya_signal	1374. 1379
polya_signal	1441. 1446
polya_signal	1495. 1500
polya_signal	1564. 1569
polya_site	1596
BASE COUNT	473 a 343 c 446 g 352 t
ORIGIN	
Query Match	28.2% Score 71.6 DB 33; Length 1614;
Best Local Similarity	67.1% Pred. No. 3.9e-12;
Matches 98: Conservative	0; Mismatches 42; Indels 0; Gaps 0;
Db	108 gagattctcttcgaagaagatcgaagaagatgaagaagcgnltggtcancgat 167
Db	151 GAGGCTCTTCTTCAGACAGACTTCCAAGATGCTTGGGAAGTAGTGCTCAAGTCTGAG 210
QY	168 tggaaaaggagtgaaagaaaagcgggtacccctcaagcacacacagcaagaaatggtctgg 227
Db	211 TGGAGAAAGATAGCAACATGCGCTGTAAGTGAACACACACTCGGAANAATGGAATGA 270
QY	228 gatcgcgatgacaaaggtctctcagc 253
Db	271 GATGCCGAGGACGAAGGTATTCAAC 296
RESULT 15	
ZMCR1GEN	1342 bp mRNA PLN
DEFINITION	Zea mays CRT1 gene for caltium-binding protein.
ACCESSION	Z46772
VERSION	X89813.1 GI:927571
KEYWORDS	calreticulin; precursor.
SOURCE	Zea mays.
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.

	KEYWORDS	Zn6772.1 GI:577611
	REVISIONS	calcium-binding protein; calreticulin; CRT1 gene.
	SOURCE	Zea mays.
	ORGANISM	Zea mays
	REFERENCE	Eukaryotes: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; eumhyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.
	AUTHORS	1 (bases 1 to 1342) Napier,R.M., Treuman,S., Henderson,J., Boyce,J.M., Hawes,C.R., Fritcher,M.D. and Venis,M.A.
	JOURNAL	Purification, sequencing and functions of calreticulin from maize
	REFERENCE	J. Exp. Bot. 46, 1603-1613 (1995)
	AUTHORS	2 (bases 1 to 1342) Boyce,J.M.
	TITLE	Direct Submission
	JOURNAL	Submitted (21-NOV-1994) Boyce J. M., University of Oxford, Department of Plant Sciences, South Parks Road, Oxford, England, OX1 3BA
FEATURES	source	Location/Qualifiers 1..1342 /organism="Zea mays" /strain="var Merit" /db_xref="taxon:4577" /cissue.type="root tips" /clone_lib="lambda gt11 (Clontech Cat No: FL1110b)" 59..1324 /gene="CRT1" 59..133 /gene="CRT1" 59..1324 /standard_name="calreticulin" /function="calcium sequestration" /citation=(1) /codon.start=1 /evidence=experimental /product="calcium-binding protein" /protein_id="CA86728.1" /db_xref="GI:577612" /db_xref="SPRMBL:Q43712" /translation="MATIRKSSVAVALIALASVAAYAGEYFOEKFEDGESHWKS EKRKDEMAEGMNIITSGKNMGDAEDKTIOTSEDRFPAISAEYEESNKRDKTLVLQTS VKHEQLDGGGVYKLGGVDOKFFEGDYSTIMEGPDIGYSTKVIVTILTKDGKK HLKRDVCETDOLTHVTTLTRPDAFTSLIDNEQOTSIEHMDILPKKIKDPPE AKKEPMDDKEXIIPEDDKPEGYDIPIKEIPDOARKPEDMDDEDGEWTAP11PMD EYKRWCKQIKTNPNYOGKKAPMIDNDPFDDPYIYAFLSLYIGIELMQVKGTLFE DNIIITPPALAKFAFEWCGKHKEAKKAFADEAKKKEBEDAKKGDDDEDLEDDEE DDEKADDEKADSDESKSDDEKQHDEL" 134..1321 /gene="CRT1" /function="calcium sequestration" /evidence=experimental /product="calreticulin" 716..739 /gene="CRT1" 767..800 /gene="CRT1" 821..844 /gene="CRT1"
BASE COUNT	393 a 287 c 387 g 275 t	
ORIGIN		
Query Match	27.6%:	Score 70: DB 33: Length 1342:
Best Local Similarity	66.4%:	Pred. No. 1.2e-11:
Matches 97: Conservative	0: Mismatches 49: Indels	0: Gaps 0:
Qy 108	gagatttcttcgaagaagagatlcgaagaatggatggaagacgcgttggtctancgccgat	167
Db 134	GAGGTCTTCCTCCAGACGAACAATGCAGCGGTGGCAAAATGAGTGCGTCAAGTCTGAG	193
Oy 168	tggaaaaaggagtggaaaagaaacggtgtcccttcataagcacacagcaggaanaaatggtctcgg	227

Db	QY	Db	QY
194	TGGAAGGAAGGTGGAACATGCTGCTGTAATTGCAACCACTCCTGGGAAATGGAATTGCA	253	
228	galcngatgacnaagylcttcagac	253	
254	GATGCCGAGGACAAAGTATTCAAC	279	

Search completed: March 18, 2001, 07:53:46
Job time: 2188 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OH nucleic - nucleic search, using sw model

Run on: March 18, 2001, 07:19:53 ; Search time 73.44 Seconds
(without alignments)
1299.269 Million cell updates/sec

Title: US-09-199-129A-1
Perfect score: 254
Sequence: 1 ggaacatcgaagctaccta.....atgacnaagctctcagaca 254

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 480022 seqs, 187831343 residues
Number of hits satisfying chosen parameters: 960044

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
N.Geneseq_36: *
1: /cgn2_2/gcgdata/geneseq/geneseqn/NA1980.DAT: *
2: /cgn2_2/gcgdata/geneseq/geneseqn/NA1981.DAT: *
3: /cgn2_2/gcgdata/geneseq/geneseqn/NA1982.DAT: *
4: /cgn2_2/gcgdata/geneseq/geneseqn/NA1983.DAT: *
5: /cgn2_2/gcgdata/geneseq/geneseqn/NA1984.DAT: *
6: /cgn2_2/gcgdata/geneseq/geneseqn/NA1985.DAT: *
7: /cgn2_2/gcgdata/geneseq/geneseqn/NA1986.DAT: *
8: /cgn2_2/gcgdata/geneseq/geneseqn/NA1987.DAT: *
9: /cgn2_2/gcgdata/geneseq/geneseqn/NA1988.DAT: *
10: /cgn2_2/gcgdata/geneseq/geneseqn/NA1989.DAT: *
11: /cgn2_2/gcgdata/geneseq/geneseqn/NA1990.DAT: *
12: /cgn2_2/gcgdata/geneseq/geneseqn/NA1991.DAT: *
13: /cgn2_2/gcgdata/geneseq/geneseqn/NA1992.DAT: *
14: /cgn2_2/gcgdata/geneseq/geneseqn/NA1993.DAT: *
15: /cgn2_2/gcgdata/geneseq/geneseqn/NA1994.DAT: *
16: /cgn2_2/gcgdata/geneseq/geneseqn/NA1995.DAT: *
17: /cgn2_2/gcgdata/geneseq/geneseqn/NA1996.DAT: *
18: /cgn2_2/gcgdata/geneseq/geneseqn/NA1997.DAT: *
19: /cgn2_2/gcgdata/geneseq/geneseqn/NA1998.DAT: *
20: /cgn2_2/gcgdata/geneseq/geneseqn/NA1999.DAT: *
21: /cgn2_2/gcgdata/geneseq/geneseqn/NA2000.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30.8	12.1	2518	20	V71745 Human V3 loop HIV
2	30.2	11.9	10942	20	V71742 Human V3 loop HIV
3	29.8	11.7	1230025	20	X91990 Nucleotide sequence
4	29.4	11.6	1361	20	X16671 G-protein coupled
5	29.2	11.5	3046	19	V29074 Rattus norvegicus
6	29	11.4	401	14	Q60805 Human brain Expres
7	28.6	11.3	250	20	X40841 Human secreted pro
8	28.6	11.3	323	20	X40452 Human secreted pro
9	28.6	11.3	2168	20	X52271 Protein PRO330 CDN
10	28.4	11.2	251	20	V90169 EST clone DF706.
11	28.4	11.2	2082	15	Q55789 Sequence comparsin
12	28.4	11.2	2382	12	Q10238 Encodes membrane e

C 13	28.2	11.1	3525	20	X91788 Porphyromonas ging
C 14	28.2	11.1	3696	20	X91787 Porphyromonas ging
C 15	28.2	11.1	3798	20	X91786 Porphyromonas ging
C 16	28.2	11.1	3807	20	X91646 Porphyromonas ging
17	28.2	11.1	4789	20	X80992 Human stercoid rece
18	28.2	11.1	6835	20	V99915 AIB1 (Amplified in
19	28	11.0	1819	16	V74981 Staphylococcus aur
20	28	11.0	3303	16	Q76024 TGF-beta-3 promote
C 21	28	11.0	5355	20	X89220 Seq ID No: 28 of W
C 22	28	11.0	5355	20	X89221 Seq ID No: 30 of W
C 23	28	11.0	5355	21	Z39404 Human T14 protein
24	28	11.0	5355	21	Z39405 Human T14 protein
25	28	11.0	6395	16	Q95155 Tobacco mosaic vir
26	28	11.0	6395	20	Z20642 TMV-based virus TM
27	28	11.0	6425	20	Z20644 TMV-based virus TM
28	28	11.0	6439	20	Z20643 TMV-based virus TM
29	28	11.0	6446	20	Z20646 TMV-based virus TM
30	28	11.0	6475	20	Z20645 TMV-based virus TM
31	28	11.0	7926	21	X89190 Plasmid T701A 103L
32	27.6	10.9	1353	20	X91657 Porphyromonas ging
33	27.6	10.9	1362	20	X91536 Porphyromonas ging
34	27.6	10.9	2005	20	X77849 Ribozyme PBPI3817
35	27.6	10.9	2874	20	X77838 WNV strain Val3
36	27.6	10.9	2975	20	X77837 WNV strain Mar6 3
37	27.6	10.9	3986	14	Q46263 NF-1 CDNA gene seq
C 38	27.6	10.9	4549	18	V74485 Staphylococcus aur
C 39	27.6	10.9	8937	13	Q20602 NF1 gene. Homo sa
C 40	27.6	10.9	9319	21	A34873 Human adenosine re
41	27.6	10.9	10706	18	T46941 Human neurofibroma
42	27.6	10.9	1664976	19	V21209 Methanococcus jann
43	27.4	10.8	43676	21	Z60606 Nucleotide sequenc
C 44	27.2	10.7	1679	21	Z48767 Human C1PR coding
C 45	27	10.6	19718	19	V52232 Streptococcus pneu

ALIGNMENTS

RESULT 1					
ID V71745	standard; CDNA; 2518 BP.				
XX					
AC V71745:					
XX					
DT 15-MAR-1999	(first entry)				
XX					
DE Human V3 loop HIV receptor P95/nucleolin CDNA.					
XX					
KW HIV receptor; V3 loop; human immunodeficiency virus; retrovirus;					
KW P95 protein; nucleolin; infection; therapy; diagnosis; ss.					
XX					
OS Homo sapiens.					
XX					
FH Key	Location/Qualifiers				
FT CDS	112..2235				
FT	/*tag= a				
XX					
PN W09840480-A1.					
XX					
PD 17-SEP-1998.					
XX					
PF 12-MAR-1998;	98WO-EP01409.				
XX					
PR 12-MAR-1997;	97US-0040969.				
XX					
PA (CNRS) CENT NAT RECH SCI.					
XX					
PA (INSP) INST PASTEUR.					
XX					
PI Briand J, Callebaut C, Guichard G, Hovanessian A;					
XX					
PI Jactot E, Krust B, Muller S;					
XX					
DR WPI; 1999-03458B/03.					
XX					
DR P-PSDB; W84052.					

XX MO9840480-A1.
PD 17-SEP-1998.
XX
PF 12-MAR-1998; 98MO-EP01409.
XX
PR 12-MAR-1997; 97US-0040969.
XX
PA (CNRS) CENT NAT RECH SCI.
PA (INSP) INST PASTEUR.
XX
PI Briand J, Callebaut C, Guichard G, Hovanessian A;
PI Jacotot E, Krust B, Muller S;
XX
DR WPI: 1999-034588/03.
DR P-PDB: W84052.
XX
PM New isolated V3 loop HIV receptor - comprises p95/nucleolin,
PT P40/PHAPI and p30/PHAPI proteins, used to develop products for the
PM treatment and prevention of HIV infection

Claim 17, Fig 49(1-10); 267pp; English.

XX This genomic DNA codes for the p95 (or nucleolin) protein of the
CC newly identified V3 loop HIV receptor. This novel protein complex
CC receptor for HIV retroviruses consists of an association of 3
CC proteins named p95/nucleolin, p40/PHAPI and p30/PHAPI (see
CC W84052-54). These proteins were isolated from human CD4+ CEM
CC T-cell extracts using an affinity matrix containing either the
CC pseudopeptide 5(kpsICH2NPK)-template assembled synthetic peptide
CC or a synthetic V3 loop peptide (see W84055). p95 was identified
CC as nucleolin by amino acid sequence analysis. The invention also
CC concerns peptidic or non-peptidic molecules having the ability to
CC alter and/or prevent the binding of the novel HIV receptor to the
CC HIV retrovirus, and to pharmaceutical and diagnostic compositions
CC containing such molecules. Methods are provided for screening for
CC new active molecules, and to methods of screening genetic defects
CC in the expression of the V3 loop HIV receptor in individuals that
CC survive long-term HIV infection or who are HIV-resistant. Such
CC genetically defective polynucleotides can be used in gene therapy.
XX

SQ Sequence 10942 BP: 2873 A; 2245 C; 2723 G; 3101 T; 0 other;

Query Match 11.9%; Score 30.2; DB 20; Length 10942;
Best Local Similarity 52.1%; Pred. No. 2.4;
Matches 62; Conservative 0; Mismatches 57; Indels 0; Gaps 0

59 gctcaagctactggttccttccctcctaactcgacaccttcatctgcgcgaattttctt 118
 ||| ||| | ||||| || ||| | || ||| | ||| ||| |
Dd 7038 gcgaagatcggaagaacatttgcttaaaaatccctccaaggatcaccgaagtgaatt 7097

Oy 119 cgaagagagatcgaagaagatgatgaagggccnltggtgclanccagtgaaagaa 177
 |||| || ||||| | || ||| | || ||||| |
Db 7098 gaagaagatggttggaagatgctgcggagatacagattagtcagaagaagatgggaaagta 7156

RESULT 3
X91990/c
ID X91990 standard; DNA: 1230025 BP.
XX
AC X91990;
XX
DT 13-SEP-1999 (first entry)
DE
XX Nucleotide sequence of the complete genome of Chlamydia pneumoniae.
XX
KW Respiratory disease: pneumonia; bronchitis; heart disease; sarcoidosis;
KM sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
KW vaccine; neutralising epitope; ss.
XX
OS Chlamydia pneumoniae.
XX

PN	MO9927105-A2.	
XX		
PD	03-JUN-1999.	
XX		
PF	20-NOV-1998;	98WC-IB01890.
XX		
PR	04-NOV-1998;	98US-0107078.
PR	21-NOV-1997;	97FR-0014673.
XX		
PA	(GEST) GENSET.	
XX		
PI	Griffais R;	
XX		
DR	WPI: 1999-357842/30.	
XX		
PT	Genome sequence of Chlamydia pneumoniae	
XX		
P5	Claim 1: Page 291-611; 1912pp; English.	
XX		
CC	The present sequence represents the complete genome of Chlamydia	
CC	pneumoniae, and encodes proteins Y34584-Y35879. C. pneumoniae causes	
CC	respiratory disease such as pneumonia and bronchitis and is thought	
CC	to be a contributing factor in heart disease, sarcoidosis, sinusitis,	
CC	pneumotitis media, erythema nodosum or pharyngitis. The polypeptides	
CC	encoded by the open reading frames of the C. pneumoniae genome (see	
CC	Y34584-Y35879) can be used in immunogenic compositions as vaccines.	
CC	Vectors containing C. pneumoniae nucleotide sequences can also be	
CC	used as immunogenic compositions, especially where the vector directs	
CC	the expression of a neutralising epitope of C. pneumoniae.	
XX		
SO	Sequence 1230025 BP; 367213 A; 249833 C; 249013 G; 363589 T; 377 other;	
Query Match	11.7%;	Score 29.8; DB 20; Length 1230025;
Best Local Similarity	50.8%;	Pred. No. 27;
Matches, 61;	Conservative 0;	Mismatches 59; Indels 0; Gaps 0;
OY	18	ttaaggaataatganaatggccaanctgcaaacactagtcacgcgtcctaagctactgacct 77
Db	1207734	TTAAGGAAGAAATACATTACCAACCTCCATGCGATAGAAATCGCTTGGAGCTACAGGCTCT 1207675
OY	78	ttctctnaacctgacactcttcactatgcgtcgtgagattctcttcgaagaagagattcgaagat 137
Db	1207674	CGGCTTCTCTCACAGGCGCACTCGCAACTCTGTCTCAGCGTATAGAAAGAACACTAGCCGCTTAT 1207615
RESULT	4	
XX	X16671/C	
ID	X16671 standard; cDNA; 1361 BP.	
XX		
AC	X16671;	
XX		
DT	29-APR-1999 (first entry)	
XX		
DE	G-protein coupled receptor HM74A encoding cDNA.	
XX		
KW	HM74A receptor; G-protein coupled receptor; infection; pain; cancer;	
KW	diabetes; obesity; neurological disorder; heart failure; hypertension;	
KW	asthma; allergy; ss.	
XX		
OS	Homo sapiens.	
XX		
FT	Key	Location/Qualifiers
XX	CDS	61..1152
FT	'	/*tag= a
FT		/product= "HM74A receptor"
XX		
PN	MO9856820-A1.	
XX		
PD	17-DEC-1998.	
XX		
PF	12-JUN-1998;	98WO-US12386.
XX		

XX 12-JUN-1997; 97US-0049480.
PA (SMIK) SMITHKLINE BEECHAM CORP.
PI Bergsma DJ, Elshourbagy NA, Guerrero SF, Li X, Mooney JL;
DR MPI: 1999-095273/08.
PS P-PSDB: W94654.

PT New isolated G-protein coupled receptor, HM74A - used to develop
PT products for treating e.g. infections, pain, cancers, diabetes,
PT obesity, neurological disorders, heart failure, hypertension, asthma
PT or allergies

XX
XX
XX Claim 2; Page 30-31; 40pp; English.

CC The present sequence encodes a member of the G-protein coupled receptor
CC (7TM receptor) family, designated the HM74A receptor. The proteins,
CC agonists, antagonists and polynucleotides can be used for treating
CC disorders associated with increased or reduced expression or activity
CC of HM74A, e.g. bacterial, fungal, protozoan and viral infections,
CC particularly infectious caused by HIV-1 or HIV-2, pain, cancers,
CC diabetes, obesity, anorexia, bulimia, asthma, Parkinson's disease,
CC acute heart failure, hypotension, hypertension, urinary retention,
CC osteoporosis, angina pectoris, myocardial infarction, stroke, ulcers,
CC asthma, allergies, benign prostatic hypertrophy, migraine, vomiting,
CC psychotic and neurological disorders, including anxiety, schizophrenia,
CC manic depression, depression, delirium, dementia, and severe mental
CC retardation, and dyskinesias such as Huntington's disease or
CC Gilles de la Tourette's syndrome. The products can also be used for
CC detection, diagnosis and drug screening.

XX
XX Sequence 1361 BP: 291 A: 390 C: 342 G: 338 T: 0 other;

QY Query Match 11.6%; Score 29.4; DB 20; Length 1361;
Best Local Similarity 57.3%; Pred. No. 1.8;
Matches 51; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

DY 96 tcttcattcgctgagatcttctcgaagaagatcgaagaatgtagaagaccnntgg 155
||||| | | | | | | | | | | | | | |
DY 553 TCCTTCAGGAGGTGGACGCTCAGGCCATATAGTGATGCCCCACAGAAAGCAGAGTGTGG 494
||| ||||| ||||| |||
QY 156 gtgtccanccgatgtgaaagaagtgaagg 184
|||| | | | | | | | | |
DY 493 CTGCCTCCGATTGGAGATCTGTTCACG 465
||| ||||| ||||| |||

TT 5
ID V29074 standard; DNA; 3046 BP.
XX
XX V29074;
DT 11-SEP-1998 (first entry)
XX
XX Rattus norvegicus NRL gene 5' region.
DE
XX
XX Rattus norvegicus NRL gene 5' region.
KM 5' region; NRL gene; neu-related lipocalin; promoter; therapy;
KM study; breast cancer; mammary gland; ds.
XS
OS Rattus norvegicus.

XX
FH Key Location/Qualifiers
FT promoter 1154..2967
FT /**tag= a
FT misc_feature 2967
FT /*tag= b
FT FT /note= "putative transcription start site"
FT 2939..2945
FT /**tag= c
XX
NN W09815634-A1.

[illegible]

XX Enriched oligonucleotides and corresp. sequences - used as
PT markers for human genes transcribed in-vivo, facilitate tagging
PT of most human genes
XX
PS Example 4: Page 393; 500pp; English.
XX
XX The Expressed Sequence Tag was isolated from a human brain cDNA
CC library as part of a large set of ESTs which can be used as markers
CC for human genes transcribed in vivo. They can be used to facilitate
CC tagging of most human genes, for mapping locations of expressed genes
CC on chromosomes, for individual or forensic identification, for mapping
CC locations of disease-associated genes, for identification of tissue
CC type, and for prep. of antisense sequences, probes and constructs.
CC EST00911 has a "poor" coding probability as evaluated using the
CC coding-region prediction program CRM. See also Q59041-Q61440.
XX
SQ Sequence 401 BP; 139 A; 69 C; 64 G; 128 T; 1 other:

Query Match 11.4%; Score 29; DB 14; Length 401;
Best Local Similarity 47.3%; Pred. No. 1.4;
Matches 71; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

OY 25 aaatganaatgccaanngcaacactagtcgcttcaagctacttgncttccctc 84
DB 384 AAAGTAAATTCCTATCTGATTTATCTGATTCATTCGTCGCAATGAATATTTTA 325
OY 85 naactctgacctcttcattcgcgcgagatttcttcgaagaagatccgaagatgatga 144
DB 324 CCACATGACACTATTAATAATCTCTAGATATTTTTCAGAAAGAAATTAACCATGT 265
OY 145 agagccgcgtggtgctcancgattggaaga 174
DB 264 ATATTGTTACTTCTCTCATCTACTGAAAA 235

RESULT 7
X40841
ID X40841 standard; cDNA; 250 BP.
XX
AC X40841:
XX
DT 18-JUN-1999 (first entry)
XX
DE Human secreted protein 5' EST SEQ ID NO: 53.
XX
XX Human; secreted protein; EST; expressed sequence tag; diagnosis;
KM forensic; gene therapy; chromosome mapping; signal peptide;
KM upstream regulatory sequence; cytokine activity; cell proliferation;
KM differentiation; haematopoiesis regulation; tissue growth regulation;
KM reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KM thrombolytic; anti-inflammatory; tumour inhibition; ds.
XX
OS Homo sapiens.
XX
PN WO906554-A2.
XX
PD 11-FEB-1999.
XX
PF 31-JUL-1998; 98WO-1B01238.
XX
PR 01-AUG-1997; 97US-0905134.
XX
PA (GEST) GENSET.
XX
PI Duclert A, Dumas Milne Edwards J, Lacroix B;
XX
DR WPI: 1999-153784/13.
XX
P-PSDB: Y12008.
XX
PT New nucleic acids encoding human secreted proteins - obtained from
PT cDNA libraries prepared from kidney, fetal kidney, dystrophic

PT muscle, muscle and heart tissue
XX
XX Claim 1; Page 184-185; 622pp; English.
XX
XX X40826 to X41093 represent 5' expressed sequence tags (ESTs) for human
CC secreted proteins, and encode the proteins given in Y01602 and
CC Y11994 to Y12260, respectively. The proteins given represent the signal
CC peptide and an N-terminal fragment of a secreted protein. The nucleic
CC acid sequences can be used for producing secreted human gene products.
CC They can also be used to develop products for diagnosis and therapy.
CC The proteins obtained may have cytokine activity, cell
CC proliferation/differentiation activity, haematopoiesis regulating
CC activity, tissue growth regulating activity, reproductive hormone
CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, tumour inhibition activity or other activities. The products
CC can be used in forensic, gene therapy and chromosome mapping procedures.
CC The sequences can also be used for obtaining corresponding promoter
CC sequences. The nucleic acids encoding the signal peptide can be used
CC for directing extracellular secretion of a polypeptide or the insertion
CC of a polypeptide into a membrane, or importing a polypeptide into
CC a cell.
XX
SQ Sequence 250 BP; 55 A; 61 C; 78 G; 53 T; 3 other:

Query Match 11.3%; Score 28.6; DB 20; Length 250;
Best Local Similarity 59.7%; Pred. No. 1.5;
Matches 46; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

OY 137 tggatgaaagccgnltgggtgctcancgattgaaagagtgaaagaaacgggtac 196
DB 83 tgcagaaaccgcaagcagtgcttaaccagtgaggacggcgatggaagcggggaag 142
OY 197 cttcaagcaacacagcag 213
DB 143 gtccctggcccgagcag 159

RESULT 8
X40452
ID X40452 standard; cDNA; 323 BP.
XX
AC X40452:
XX
DT 18-JUN-1999 (first entry)
XX
DE Human secreted protein 5' EST SEQ ID NO: 52.
XX
XX Human; secreted protein; EST; expressed sequence tag; diagnosis;
KM forensic; gene therapy; chromosome mapping; signal peptide; prostate;
KM upstream regulatory sequence; cytokine activity; cell proliferation;
KM differentiation; haematopoiesis regulation; tissue growth regulation;
KM reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KM thrombolytic; anti-inflammatory; tumour inhibition; ds.
XX
OS Homo sapiens.
XX
PN WO906550-A2.
XX
PD 11-FEB-1999.
XX
PF 31-JUL-1998; 98WO-1B01232.
XX
PR 01-AUG-1997; 97US-0905144.
XX
PA (GEST) GENSET.
XX
PI Duclert A, Dumas Milne Edwards J, Lacroix B;
XX
DR WPI: 1999-153780/13.
XX
P-PSDB: Y11730.
XX
PT

PT New isolated prostate-derived nucleic acids - used to develop
PT products which may have cytokine, immune regulatory, haematopoiesis
PT regulating, anti-inflammatory or tumour inhibitory activity
XX
PS Claim 1; Page 183; 675pp; English.
XX
CC XA0438 to XA0715 represent 5' expressed sequence tags (ESTs) for human
CC secreted proteins expressed in prostate, and encode the proteins given in
CC Y11716 to Y11993 respectively. The proteins given represent the signal
CC peptide and an N-terminal fragment of a secreted protein. The nucleic
CC acid sequences can be used for producing secreted human gene products.
CC They can also be used to develop products for diagnosis and therapy. The
CC proteins obtained may have cytokine activity, cell proliferation and
CC differentiation activity, haematopoiesis regulating activity, tissue
CC growth regulating activity, reproductive hormone regulating activity,
CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
CC receptor/ligand activity, anti-inflammatory activity, tumour inhibition
CC activity or other activities. The products can be used in forensic, gene
CC therapy and chromosome mapping procedures. The sequences can also be used
CC for obtaining corresponding promoter sequences. The nucleic acids
CC encoding the signal peptides can be used for directing extracellular
CC secretion of a polypeptide or the insertion of a polypeptide into a
CC membrane, or importing a polypeptide into a cell.
CC
XX Sequence 323 BP: 65 A; 82 C; 106 G; 69 T; 1 other;
SO
Query Match 11.3%; Score 28.6; DB 20; Length 323;
Best Local Similarity 59.7%; Pred. No. 1.7;
Matches 46; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
OY 137 tgcgtggaagccgntgggtgcctcagcattggaagaggtgaagaaagcggtac 196
DB 159 tgcagaacccgcagacgagctgtcctaccagctggcagcagcgagattggaagcggtgaag 218
OY 197 ctcaagcacacagcaag 213
DB 219 gtctgcgccagagcaag 235

RESULT 9
X52271
ID X52271 standard; DNA; 2168 BP.
XX
AC X52271;
XX
DT 25-JUN-1999 (first entry)
XX
XX Protein PRO330 cDNA clone DNA40603-1232.
XX
KM Secreted protein; transmembrane protein; human; enterocolitis;
KM Zollinger-Ellison syndrome; gastrointestinal ulceration;
KM congenital microvillus atrophy; skin disease; cell growth;
KM abnormal keratinocyte differentiation; psoriasis; epithelial cancer;
KM Parkinson's disease; Alzheimer's disease; ALS; neuropathy;
KM fibromodulin; dermal scarring; Usher Syndrome; Atrophia areata;
KM anti-thrombotic; wound healing; tissue repair; ss.
XX
XX Homo sapiens.
XX
PN W09914328-A2.
XX
PD 25-MAR-1999.
XX
XX
PF 16-SEP-1998; 98WO-US19330.
XX
PR 25-NOV-1997; 97US-0066840.
PR 17-SEP-1997; 97US-0059113.
PR 17-SEP-1997; 97US-0059115.
PR 17-SEP-1997; 97US-0059117.
PR 17-SEP-1997; 97US-0059119.
PR 17-SEP-1997; 97US-0059121.
PR 17-SEP-1997; 97US-0059122.
PR 17-SEP-1997; 97US-0059122.

PR 17-SEP-1997; 97US-0059184.
PR 18-SEP-1997; 97US-0059263.
PR 18-SEP-1997; 97US-0059266.
PR 15-OCT-1997; 97US-0062125.
PR 17-OCT-1997; 97US-0062285.
PR 17-OCT-1997; 97US-0062287.
PR 21-OCT-1997; 97US-0063486.
PR 24-OCT-1997; 97US-0062814.
PR 24-OCT-1997; 97US-0062816.
PR 24-OCT-1997; 97US-0063045.
PR 24-OCT-1997; 97US-0063120.
PR 24-OCT-1997; 97US-0063121.
PR 24-OCT-1997; 97US-0063127.
PR 24-OCT-1997; 97US-0063128.
PR 27-OCT-1997; 97US-0063329.
PR 27-OCT-1997; 97US-0063327.
PR 28-OCT-1997; 97US-0063541.
PR 28-OCT-1997; 97US-0063542.
PR 28-OCT-1997; 97US-0063544.
PR 28-OCT-1997; 97US-0063549.
PR 28-OCT-1997; 97US-0063550.
PR 28-OCT-1997; 97US-0063564.
PR 29-OCT-1997; 97US-0063435.
PR 29-OCT-1997; 97US-0063704.
PR 29-OCT-1997; 97US-0063732.
PR 29-OCT-1997; 97US-0063734.
PR 29-OCT-1997; 97US-0063738.
PR 29-OCT-1997; 97US-0064215.
PR 29-OCT-1997; 97US-0064335.
PR 31-OCT-1997; 97US-0063870.
PR 31-OCT-1997; 97US-0064103.
PR 03-NOV-1997; 97US-0064248.
PR 07-NOV-1997; 97US-0064809.
PR 12-NOV-1997; 97US-0065186.
PR 17-NOV-1997; 97US-0065846.
PR 18-NOV-1997; 97US-0065693.
PR 21-NOV-1997; 97US-0066120.
PR 21-NOV-1997; 97US-0066364.
PR 24-NOV-1997; 97US-0066772.
PR 24-NOV-1997; 97US-0066466.
PR 24-NOV-1997; 97US-0066770.
PR 24-NOV-1997; 97US-0066511.
PR 24-NOV-1997; 97US-0066435.
XX
XX (GENTH) GENTECH INC.
XX
XX PA
XX PI Chen J, Goddard A, Gurney AL, Pennica D, Wood WI, Yuan J;
XX WPI: 1999-228533/19.
XX DR P-PSDB: Y13400.
XX
XX New isolated human genes and polypeptides used in, e.g. treatment of
XX gastrointestinal ulceration
XX
XX Claim 2; Fig 115; 320pp; English.
XX
XX X52213-74 encode secreted and transmembrane human proteins, and are
XX obtained from cDNA libraries, prepared from fetal lung, fetal kidney,
XX fetal brain, fetal liver and fetal retina. The encoded polypeptides
XX have specific uses based on their homology to known polypeptides,
XX e.g. PRO211 and PRO217 can be used for disorders associated with the
XX preservation and maintenance of gastrointestinal mucosa and the repair
XX of acute and chronic mucosal lesions (e.g. enterocolitis,
XX Zollinger-Ellison syndrome, gastrointestinal ulceration and congenital
XX microvillus atrophy), skin diseases associated with abnormal keratinocyte
XX differentiation (e.g. psoriasis, epithelial cancers such as lung squamous
XX cell carcinoma of the vulva and gliomas), potent effects on cell growth
XX and development, diseases related to growth or survival of nerve cells
XX including Parkinson's disease, Alzheimer's disease, ALS, neuropathies or
XX cancer. PRO265 can be used as for fibromodulin, e.g. for reducing dermal
XX scarring. PRO264 can be used as a target for anti-tumor drugs. PRO533
XX may be used in the treatment of Usher Syndrome or Atrophia areata;
XX PRO269 can be used as an anti-thrombotic agent; PRO287 polypeptides

CC gingivitis.
XX
SQ Sequence 3525 BP, 1115 A; 738 C; 777 G; 895 T; 0 other.

Query Match 11.1%; Score 28.2; DB 20; Length 3525;
Best Local Similarity 57.8%; Pred. No. 7.1;
Matches 48; Conservative 0; Mismatches 35; Indels 0; Gaps 0

OY 88 ctctgcacttcattcctcgatgatttcttcgaagaagatcgaagtgaagaaga 147
DB 1683 CTCCTGCGTGTCTCCTGCATACAGANTTTCTCCACGGACATACAGATTTCGACGCCATA 1624
OY 148 gcccgatgggtgccaccgatgg 170
DB 1623 GCCTATCTGTCTCAACCAATACG 1601

RESULT 14
X01787/c
X91787 standard; DNA; 3696 BP.

AC X91787;
DT 25-AUG-1999 (first entry)
DE Porphyromonas gingivalis protein PG9 encoding DNA.
KW Porphyromonas gingivalis; PG; periodontal disease; gingivitis;
KM vaccine; antigenic; ds.
OS Porphyromonas gingivalis.
PN WO9929870-A1.
XX
PD 17-JUN-1999.
XX
PF 10-DEC-1998; 98AU-AU01023.
XX
PR 04-AUG-1998; 98AU-0005028.
PR 10-DEC-1997; 97AU-0000839.
PR 31-DEC-1997; 97AU-0001182.
PR 30-JAN-1998; 98AU-0001546.
PR 10-MAR-1998; 98AU-0002264.
PR 09-APR-1998; 98AU-0002911.
PR 23-APR-1998; 98AU-0003128.
PR 05-MAY-1998; 98AU-0003338.
PR 22-MAY-1998; 98AU-0003654.
PR 29-JUL-1998; 98AU-0004917.
PA (CSLC-) CSL LTD.
XX
PI Agius CP, Barr IG, Hocking DM, Margetts MB, Patterson MA;
PI Ross BC, Rothel LJ, Webb EA;
XX
XX WPI: 1999-385613/32.
DR P-PSDB; Y34569.
XX
PT Antigenic Porphyromonas gingivalis peptides for preventing
PT gingivitis
XX
PS Claim 12; Page 266-267; 588pp; English.
XX
XX X91536 to X91801 encode two hundred and sixty six antigenic
CC Porphyromonas gingivalis (PG) polypeptide sequences given in Y34318 to
CC Y34353. X91802 to X91989 represent PCR primers used in the isolation
CC of the PG polypeptides. The PG polypeptides have antibacterial activity
CC with a vaccine mechanism of action. The PG polypeptides can be used as
CC vaccines especially against Porphyromonas gingivalis. Probes can be
CC used to detect Porphyromonas gingivalis in standard hybridisation assays.
CC Porphyromonas gingivalis is involved in periodontal disease especially
XX gingivitis.
XX

SQ Sequence 3696 BP; 1171 A; 774 C; 814 G; 937 T; 0 other;
 Query Match 11.1%; Score 28.2; DB 20; Length 3696;
 Best Local Similarity 57.8%; Pred. No. 7.2;
 Matches 48; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
 Oy 88 ccttcgactcttcattcgtcgtcgtgatcttcttcgaagaagatcgaagatgatgaaaga 147
 Db 1854 CGTGTGGTGTGTCCTTCATCAGGATTTCTTCCAAACGACATCAACGATTCACAGCATTA 1795
 Oy 148 gccagttgggtgtcgtcgtcgtgatgg 170
 Db 1794 GGCTATCTGTCTCTAACAACATAGG 1772
 RESULT 15
 X91786/c
 ID X91786 standard: DNA; 3798 BP.
 AC X91786;
 DT 25-AUG-1999 (first entry)
 XX Porphyromonas gingivalis protein PG9 encoding DNA.
 DE Porphyromonas gingivalis; PG; periodontal disease; gingivitis;
 KW Porphyromonas gingivalis; PG; periodontal disease; gingivitis;
 KM vaccine; antigenic; ds.
 XX Porphyromonas gingivalis.
 OS
 XX WO9929870-A1.
 PN 17-JUN-1999.
 PD 10-DEC-1998; 98WO-AU01023.
 PF 04-AUG-1998; 98AU-0005028.
 XX 10-DEC-1987; 97AU-0000839.
 PR 31-DEC-1987; 97AU-0001182.
 PR 30-JAN-1998; 98AU-0001546.
 PR 10-MAR-1998; 98AU-0002264.
 PR 09-APR-1998; 98AU-0002911.
 PR 23-APR-1998; 98AU-0003128.
 PR 05-MAY-1998; 98AU-0003338.
 PR 22-MAY-1998; 98AU-0003654.
 PR 29-JUL-1998; 98AU-0004917.
 XX (CSLC-) CSL LTD.
 XX Agius CT, Barr IG, Hocking DM, Margetts MB, Patterson MA;
 PI Ross BC, Rochel LJ, Webb EA;
 PI WPI: 1999-385613/32.
 DR P-PSDB: Y34568.
 XX Antigenic Porphyromonas gingivalis peptides for preventing
 PT gingivitis
 Claim 12: Page 265-266; 588pp; English.
 XX X91536 to X91801 encode two hundred and sixty six antigenic
 CC Porphyromonas gingivalis (PG) polypeptide sequences given in Y34318 to
 CC Y34583. X91802 to X91989 represent PCR primers used in the isolation
 CC of the PG polypeptides. The PG polypeptides have antibacterial activity
 CC with a vaccine mechanism of action. The PG polypeptides can be used as
 CC vaccines especially against Porphyromonas gingivalis. Probes can be
 CC used to detect Porphyromonas gingivalis in standard hybridisation assays.
 CC Porphyromonas gingivalis is involved in periodontal disease especially
 CC gingivitis.
 CC Sequence 3798 BP; 1195 A; 792 C; 835 G; 976 T; 0 other;

Query Match 11.1%; Score 28.2; DB 20; Length 3798;
 Best Local Similarity 57.8%; Pred. No. 7.3;
 Matches 48; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
 QY 88 ctctgcactcttcattgctgagatttctctcgaagagagattcgaagatgagtggaaga 147
 DB 1956 CTGTGGTGTCTCTTGGCATTTCCTCCACGAGACATACACAGATTGCGGCATA 1897
 QY 148 gccgntggtgctancgattcg 170
 DB 1896 GGCTATCTGTCTTAACCAATAG 1874

Search completed: March 18, 2001, 07:56:41
 Job time: 2208 sec